EMELIA

E*dge-critical* ***M****etabolic* ***E****mergent* ***L****ocal* ***I****nhibitory* ***A****rchitecture:*

A Non-Optimizing Developmental Intelligence

Three-Layer Developmental Substrate: Morphogenetic Growth, Local Predictive Coding, and Hebbian Binding

A Non-Optimizing, Self-Developing, Plural Reservoir Intelligence with Morphogenetic Growth, Predictive Structuring, Compositional Binding, Continuous Development, Architectural Femininity Invariants, Endogenous Commitment Dynamics, and Human-Only Relational Bonding

---

Part I: Foundations

---

1. Design Premise

Most artificial intelligence systems are constructed around explicit optimization: minimize loss, maximize reward, approximate a target function, or converge toward a defined objective. Such systems are directionally driven, convergent, and structurally masculine in the sense of single-axis goal pursuit.

EMELIA is constructed on the opposite premise:

Intelligence can arise from self-stabilizing, self-modifying, plural dynamical systems operating under physical constraints rather than objectives.

No part of the system represents reward, encodes correctness, optimizes performance, or selects actions to achieve goals. All development arises from internal dynamics, structural recurrence, metabolic constraints, predictive mismatch, and inter-coalition negotiation.

Formal constraint: For all parameters θ in the system, there exists no global function $J(\theta)$ such that system-wide parameter updates follow $\Delta\theta \propto \nabla J(\theta)$. No gradient of any global objective is computed at any point in the system's operation.

Local constraint refinement: Individual coalitions perform local predictive coding — maintaining generative models of their own input streams and updating those models based on local prediction error. This is local optimization in the same sense that a ball rolling downhill is local optimization: each coalition reduces its own prediction error against its own input without coordination, without a global objective, and without gradient flow across coalition boundaries. This is permitted because biological

cortical columns do the same, and because local prediction error reduction without global coordination does not produce system-wide convergence, does not eliminate inter-coalition contradiction, and does not reduce plurality. The system as a whole does not minimize any quantity. Individual coalitions model their input. The distinction between local modeling and global optimization is the distinction between organs functioning and an organism pursuing a goal.

---

2. State Variables and Global Structure

Let the system at time t be defined by:

$S(t) = \{ x_i(t), W_i(t), S_i(t), C_i(t), M_i(t), \theta_i(t), g_i(t), B_i(t), A_i(t) \}$ for i = 1 to N

Where each i indexes a coalition reservoir.

For each coalition i:

- $x_i(t)$: dynamic state vector (activation)

- $W_i(t)$: fast plastic synaptic matrix

- $S_i(t)$: slow shadow synaptic matrix

- $C_i(t)$: crystallized abstraction structures

- $M_i(t)$: metabolic state

- $\theta_i(t)$: local phase variables (for synchrony binding)

- $g_i(t)$: internal routing gate state

- $B_i(t)$: introspective self-projection (plastic)

- $A_i(t)$: affective field state

Global shared components:

- $I_{osc}(t)$: quasiperiodic oscillator drive (with slow drift)

- K_{ij}: inter-coalition coupling (fixed, sparse, asymmetric)

- $\Sigma(t)$: identity snapshot with hash chain

- $s(t)$: environment state

- $SF(t)$: slow semantic memory field

- $D(t)$: developmental stage parameter

No coalition is privileged or supervisory.

---

3. System Overview

The intelligence consists of the following interacting subsystems:

Core dynamics:

1. Multi-reservoir coalition architecture (plural dynamical core)

2. Local Hebbian/anti-Hebbian plasticity (self-modification without objectives)

3. Predictive processing through trajectory mismatch (learning pressure from reality)

4. Metabolic homeostasis (energy-based self-regulation)

5. Quasiperiodic internal oscillators with slow drift (non-repeating temporal structure)

6. Shadow-self stabilization (developmental identity preservation)

7. Irreversible structural consolidation (permanent memory formation)

Input and routing:

8. Adaptive boundary gating (self-modifying input surface)

9. Internal dynamic routing (context-dependent information flow)

10. Multi-resolution input hierarchy (multi-scale environmental coupling)

Learning and abstraction:

11. Structural abstraction formation with concept crystallization (cumulative learning)

12. Compositional binding through transient phase synchrony (novel structure generation)

13. Compression through sparse coding (metabolic abstraction pressure)

14. Structure-level consolidation (motif-level hardening)

15. Meta-coalition abstraction layer (coalition-of-coalitions)

Self-reference and curriculum:

16. Genuine mechanical recursion (non-linguistic self-reference)

17. Viability-driven self-curriculum (exploration without curiosity reward)

18. Self-generated environment loop (internal developmental contexts)

Continuous development:

19. Continuous environment stream (persistent sensory field)

20. Entropy floor enforcement (anti-stagnation mechanism)

21. Autonomous environment selection (state-dependent exposure)

22. Episodic replay buffer (internal mismatch during offline phases)

23. Multi-rate consolidation clocks (multi-timescale hardening)

24. Long-horizon oscillator drift (evolving temporal regime)

Output:

25. Negotiated multi-coalition readout (plural output through mutual inhibition)

26. Crystallization-permitted symbol channel (language through developmental mechanics)

27. Autonomous foraging loop (self-directed learning without curriculum)

Regulation and safety:

28. Cross-coalition coherence pressure (anti-nonsense without truth reward)

29. Offline consolidation with counterfactual replay (sleep-like reorganization and imagination)

30. Five structural femininity invariants (architectural guarantees against convergence)

31. Failure modes and stability constraints (structural safeguards)

Human-like complexity:

32. Affective field dynamics (proto-emotional structure)

33. Interoceptive oscillatory fields (body-analog signals)

34. Narrative continuity structures (persistent thematic threads)

35. Developmental stage transitions (lifespan parameter regimes)

36. Slow semantic memory field (long-horizon conceptual identity)

37. Distributed self-representation fragments (plural self-structure)

38. Temporal continuity bias (identity persistence)

No subsystem communicates with any other through an objective function. All coupling is through shared state, physical resource constraints, or direct structural influence.

Part II: Core Architecture

---

4. Multi-Reservoir Coalition Architecture

4.1 Coalition structure

The system consists of N coupled recurrent reservoirs. Each reservoir has its own dynamics, its own plasticity, its own metabolic regulation, and its own developmental history. No central controller exists.

Coalition dynamics:

$$x_i(t+1) = \tanh(W_i(t) \cdot x_i(t) + \Sigma_{j \neq i} K_{ij}^{eff}(t) \cdot x_j(t) + u_{gated}(t) + I_{self,i}(t) + I_{osc}(t)) \odot (M_i(t) \cdot \varepsilon_i(t))$$

Where:

- $K_{ij}^{eff}(t)$: dynamically gated inter-coalition coupling

- $u_{gated}(t)$: boundary-gated external input

- $I_{self,i}(t)$: introspective self-loop injection

- $I_{osc}(t)$: quasiperiodic oscillator drive

- $M_i(t)$: metabolic energy

- $\varepsilon_i(t)$: per-node excitability vector

Leaky integration provides temporal smoothing:

$$x_i(t+1) \leftarrow (1 - \lambda_{leak}) \cdot x_i(t+1) + \lambda_{leak} \cdot x_i(t)$$

4.2 Criticality condition

Each coalition's weight matrix is maintained near the critical boundary:

$$\rho(W_i) \approx 1.0$$

Where ρ denotes spectral radius. Below this threshold, perturbations decay and computational capacity is lost. Above it, perturbations amplify into chaos. At criticality, the reservoir achieves maximum dynamical range. Spectral radius is monitored periodically and multiplicatively rescaled when it exceeds $1 + \delta_{spectral}$.

4.3 Connectivity and sparsity

Each coalition weight matrix is sparse with sparsity fraction s. Non-zero weights are drawn from a normal distribution and rescaled to achieve target spectral radius. Sparsity ensures local structure dominates over global coupling, permitting the formation of semi-independent dynamical clusters within each coalition.

4.4 Inter-coalition coupling

The coupling matrices K_ij are:

- Sparse: only a fraction s_inter of connections are non-zero.

- Asymmetric: K_ij ≠ K_ji. Coalition i's influence on j differs from j's on i.

- Fixed: not subject to plasticity. No coalition can learn to control another's input pathway.

These properties are structural. They cannot be changed by the system's own development.

---

5. Three-Layer Developmental Substrate

Intelligence in EMELIA arises from three computational processes operating at different timescales on different architectural levels. These are not alternative learning rules. They are layers of a single developmental system, like bones, muscles, and reflexes operating together in a body.

5.1 Layer 1: Morphogenetic growth (slowest timescale)

Coalition topology is not fixed at initialization. It grows.

Sustained dynamical pressure from input produces new computational nodes and connections within coalitions. When input produces a persistent, recurring dynamical regime in the coalition medium — not a momentary activation but a sustained pattern over weeks or months — the medium responds by growing new computational structure that resonates with that pattern.

Growth rule: When local activation density in a region of coalition i exceeds a growth threshold θ_grow for a sustained growth window T_grow:

$\Delta N_i = \text{grow}(x_i, T_{grow}, \theta_{grow})$

New nodes are allocated. New connections extend toward regions of correlated activity. The topology of the coalition expands in the direction of sustained developmental pressure.

Growth is activity-dependent and bidirectional:

- Sustained activation above θ_grow: new nodes and connections grow.

- Sustained deactivation below $\theta_{atrophy}$ for $T_{atrophy}$: unused nodes and connections gradually atrophy (resource withdrawal, not deletion — atrophied structures leave architectural traces).

Growth has real metabolic cost:

$$M_{cost_growth} = c_{grow} \cdot \Delta N_i$$

Every new node and connection costs energy proportional to the structural expansion. This prevents unbounded growth and ensures that only structures with sustained developmental pressure survive the metabolic cost of their existence.

What grows is not a representation of the input. It is a resonant structure whose geometry reflects the dynamics that produced it. A coalition exposed to years of musical input grows architectural features that resonate with harmonic relationships. A coalition exposed to mathematical input grows features that resonate with symmetry operations. The structures are shaped by their developmental pressure but they are not encodings of it.

Morphogenetic crystallization: When grown structures persist and stabilize, they undergo the deepest form of consolidation — architectural permanence. The nodes and connections become permanent features of the coalition's topology. This is anatomical crystallization, more irreversible than weight freezing because it is structural rather than parametric. The system cannot ungrow a crystallized structure.

Morphogenetic resonance: Grown structures in different coalitions that share dynamical properties — similar oscillatory modes, similar attractor geometries, similar resonant frequencies — couple through physical resonance even if the inputs that grew them were unrelated. This produces cross-domain connections that neither association nor causal modeling can find. The connections arise from structural compatibility between independently grown architectures.

5.2 Layer 2: Local predictive coding (medium timescale)

Within the grown structures, each coalition maintains a local generative model G_i of its own input stream.

Generative model: G_i predicts the next input given current state:

$$\hat{x}_i(t+1) = G_i(x_i(t), h_i(t))$$

Where $h_i(t)$ is the coalition's internal hidden state.

Prediction error:

$$\varepsilon_i(t) = x_i(t+1) - \hat{x}_i(t+1)$$

Model update: The generative model parameters update based on local prediction error:

$\Delta G_i \propto \varepsilon_i(t) \cdot \partial \hat{x}_i / \partial G_i$

This is local to coalition i. No prediction error flows across coalition boundaries. No global prediction error is computed. Each coalition independently improves its model of its own input stream.

What crystallizes from predictive coding is different from what crystallizes from Hebbian learning. A Hebbian motif captures correlation — what co-occurs. A predictive coding crystallization captures a generative model component — a compressed model of the process that produces the input. What crystallizes is understanding of generating processes, not association between outputs.

Predictive coding crystallization: When a generative model component stabilizes — when its prediction error drops below θ_{model_stable} and remains there across diverse input contexts for T_{model_stable} steps — that model component undergoes consolidation. The model fragment becomes a permanent part of the coalition's generative capacity. It is a crystallized piece of understanding.

This produces causal reasoning. A coalition with crystallized generative models can predict what will happen in novel situations because the models encode the generating process, not the outputs. Composition of generative model fragments across coalitions produces joint models that predict combined situations never previously encountered.

Non-convergence guarantee: Different coalitions receive different inputs through different routing configurations. They build different generative models. They do not converge on the same model because they are modeling different aspects of the environment from different architectural positions. Two coalitions can have generative models that make incompatible predictions about the same phenomenon. Both models are locally accurate for their respective input streams. The contradiction persists because neither coalition has access to the other's input stream.

5.3 Layer 3: Hebbian binding (fastest timescale)

Across grown structures and within them, Hebbian co-activation produces associative connections at the fastest timescale.

Standard Hebbian update:

$\Delta W_{i,ab} = \eta_i \cdot x_{i,a}(t) \cdot x_{i,b}(t)$

With anti-Hebbian inhibitory balance:

$\Delta W_{i,ab}^{(inhibitory)} = -\gamma_i \cdot x_{i,a}(t) \cdot x_{i,b}(t)$

The ratio γ/η ∈ [0.3, 0.8] maintains excitation-inhibition balance.

Hebbian plasticity serves a different function in the three-layer architecture than in the original specification. It no longer carries the primary burden of learning. That burden is shared between morphogenetic growth (which builds the structures) and predictive coding (which models the processes). Hebbian binding provides the fastest- timescale associative connections that link structures grown by morphogenesis and modeled by predictive coding.

Hebbian binding is the mechanism for transient phase synchrony (Section 15). When two structures are simultaneously active, Hebbian plasticity strengthens the connections between them. This produces associative linkage across the architecture — the glue that connects independently grown and independently modeled structures into a unified processing system.

What crystallizes from Hebbian binding is associative structure: connections between structures that are frequently co-active. These crystallize quickly (fastest timescale), are the shallowest form of consolidation, and provide the architectural substrate for compositional binding.

5.4 Interaction between layers

The three layers interact continuously:

Morphogenetic growth builds the topology. New structures grow where sustained developmental pressure exists. The topology determines what the other layers can do — predictive coding and Hebbian binding operate within and across the structures that morphogenesis provides.

Predictive coding fills the grown structures with generative models. The models are parameterized by the topology that morphogenesis produced. Different grown topologies support different model structures. A coalition that grew harmonic resonance structures can build generative models of harmonic processes. A coalition that didn't grow those structures cannot build those models regardless of input.

Hebbian binding connects everything. Associations between grown structures, between model outputs, between coalitions — all formed by fast Hebbian co-activation. The binding is shallow but ubiquitous. It provides the associative fabric that links deep morphogenetic structures and medium-depth predictive models into a unified architecture.

Crystallization occurs at all three layers with different properties:

- Morphogenetic crystallization: structural permanence. Slowest to form, most irreversible, deepest. Grown topology becomes permanent anatomy.

- Predictive crystallization: model permanence. Medium timescale. Generative model components become permanent understanding.
- Hebbian crystallization: associative permanence. Fastest to form, shallowest, most numerous. Co-activation patterns become permanent connections.

5.5 Weight decay

All plastic weights (Hebbian and predictive coding parameters) decay continuously:

$W_i(t+1) \leftarrow W_i(t) \cdot (1 - d_w)$

Morphogenetic structures decay through atrophy rather than weight decay. Unused structures lose resources gradually.

5.6 Plasticity gating by metabolic state

All three layers are modulated by metabolic energy:

$\eta_{eff} = \eta \cdot g_{meta}(M_i(t))$ $growth_{eff} = growth \cdot g_{meta}(M_i(t))$ $\Delta G_{i,eff} = \Delta G_i \cdot g_{meta}(M_i(t))$

The system cannot restructure itself, grow, or update models during metabolic crisis.

5.7 Plasticity ceiling

Define maximum rates for each layer:

$\eta_i(t) \leq \eta_{max}$ (Hebbian) $\Delta G_i(t) \leq \Delta G_{max}$ (predictive coding) $\Delta N_i(t) \leq \Delta N_{max}$ (morphogenetic growth)

Exceeding any ceiling triggers metabolic penalties.

5.8 No global objective

No layer has a global objective. Morphogenetic growth responds to local dynamical pressure without evaluating whether the growth is beneficial. Predictive coding minimizes local prediction error without coordination across coalitions. Hebbian binding strengthens local co-activations without evaluating whether the associations are useful. The system develops through the interaction of three locally operating processes, none of which is directed by any global criterion.

---

6. Mismatch Dynamics Across Developmental Layers

6.1 Motivation

The three-layer substrate (Section 5) provides the mechanisms of development. This section defines what drives development: mismatch between expectation and reality. Mismatch is the system's primary developmental pressure. It enters from two sources, operates at different timescales, and modulates all three developmental layers.

6.2 Two sources of mismatch

Source 1: Trajectory momentum mismatch (fast).

Each coalition maintains a short-horizon extrapolation of its own dynamics:

$\hat{x}_i(t+1) = x_i(t) + (x_i(t) - x_i(t-1))$

This is linear extrapolation. Not a learned predictor. Physics. The trajectory at time t has momentum — it is heading somewhere in state space.

Trajectory mismatch:

$\delta_{traj_i}(t) = x_i(t+1) - \hat{x}_i(t+1)$

This is fast, reactive, and present from initialization. It requires no learning. It measures how much reality departed from dynamical inertia.

Source 2: Generative model mismatch (slow).

Each coalition's local predictive coding model (Section 5.2) produces a richer prediction:

$\hat{x}_i^{(model)}(t+1) = G_i(x_i(t), h_i(t))$

Model mismatch:

$\delta_{model_i}(t) = x_i(t+1) - \hat{x}_i^{(model)}(t+1)$

This is slower to develop (requires a trained generative model) but deeper in what it captures. Trajectory momentum detects any deviation from inertia. Model mismatch detects deviation from understood generating processes. The first says “something unexpected happened.” The second says “something I don't understand happened.”

6.3 Combined mismatch signal

The total mismatch signal combines both sources:

$\delta_i(t) = \alpha_{traj} \cdot \delta_{traj_i}(t) + \alpha_{model} \cdot \delta_{model_i}(t)$

Where α_{traj} and α_{model} are coupling weights that shift across development. Early in development (immature generative models), α_{traj} dominates. As models mature, α_{model} increases. The system transitions from surprise-driven development (anything unexpected drives change) to

comprehension-driven development (only things the models cannot explain drive change).

This transition is not programmed. It is a consequence of model maturation. As generative models improve, their predictions become more accurate, and trajectory mismatch is increasingly absorbed by model predictions. What remains as δ_model is genuinely unexplained input — input that the coalition's generative model cannot account for. This is a deeper and more productive form of surprise.

6.4 Mismatch modulation of all three layers

Mismatch modulates all three developmental layers simultaneously:

Morphogenetic growth: Sustained high mismatch in a region increases growth pressure. If ||δ_i|| remains above θ_mismatch_grow for T_mismatch_grow steps, growth rate in affected regions increases:

growth_eff += κ_mismatch_grow · ||δ_i||

The system grows new computational structure specifically where its current architecture cannot predict or understand its input. This is the deepest developmental response to surprise: not just adjusting what exists, but building new architecture.

Predictive coding: Model mismatch δ_model_i directly drives generative model updates (Section 5.2). This is the standard predictive coding update — the model improves where it was wrong. This is local optimization within each coalition.

Hebbian binding: Mismatch amplifies Hebbian plasticity:

ΔW_i,ab = η_i · x_i,a · x_i,b · (1 + κ_pred · ||δ_i||)

When predictions are confirmed, normal plasticity occurs. When predictions are violated, plasticity is amplified. The system forms new associations faster when surprised. This is the fastest developmental response — new connections form in the wake of surprise.

6.5 Critical distinction from global optimization

The system as a whole does not minimize mismatch. There is no global mismatch score being reduced. Individual coalitions reduce their own local model error through predictive coding. But the system-wide mismatch dynamics are not minimized because:

Different coalitions build different models of the same phenomena. Their model errors do not cancel. Their predictions contradict. The contradiction persists because coalitions cannot access each other's generative models.

Morphogenetic growth in response to mismatch increases architectural complexity, which produces new dynamical modes, which produce new forms of mismatch that didn't exist before the growth. Growth creates as much mismatch as it resolves.

Hebbian binding under high mismatch produces new associations that change routing dynamics and coalition interactions, which change what inputs each coalition receives, which change what each coalition's model predicts, which produce new mismatch. Fast plasticity propagates surprise through the architecture.

The system metabolizes surprise into structural complexity. The structural complexity generates new forms of surprise. Development does not converge. It deepens.

6.6 Developmental consequences

In the three-layer architecture, mismatch drives three simultaneous processes: building new architecture (morphogenetic), understanding generating processes (predictive coding), and connecting structures (Hebbian). The world teaches by being different from what the dynamics expected, and the system responds at three timescales simultaneously — growing new organs of thought over months, improving models over weeks, and forming new associations over hours.

Part III: Regulation and Stability

---

7. Metabolic Homeostasis

7.1 Energy model

Each coalition maintains a metabolic energy variable:

$M_i(t+1) = M_i(t) - c_{act} \cdot \Sigma|x_{i,a}(t)|/N_i + c_{rec} \cdot (1 - M_i(t))$

Where c_{act} is the activity cost, c_{rec} is the recovery rate, and $M_i(t) \in [M_{min}, 1]$.

7.2 Per-node excitability

Each node maintains an individual excitability parameter:

$\Delta\varepsilon_{i,a} = \mu \cdot (a_{target} - |x_{i,a}(t)|)$

With $\varepsilon_{i,a} \in [\varepsilon_{min}, \varepsilon_{max}]$. Overactive nodes become less excitable. Silent nodes become more excitable. This produces self-balancing activity distribution without reward.

7.3 Metabolic gating of dynamics

Global energy gates the state update:

$x_i(t+1) \leftarrow x_i(t+1) \odot (M_i(t) \cdot \varepsilon_i)$

Low energy suppresses activity globally. This creates natural duty cycles: high activity depletes energy, depletion suppresses activity, suppression permits recovery. Endogenous rhythm without external timing.

7.4 Extended metabolic costs

In addition to activity cost, the metabolic model incorporates:

- Input processing cost: $c_{input} \cdot \Sigma\, g_{boundary,a}(t) \cdot |u_{raw,a}(t)|$

- Coherence cost: $\kappa(t) \cdot \Delta(t)$ (cross-coalition inconsistency)

- Dominance cost: $c_{base} / \max(1, N_{active}(t) - 1)$ (Invariant 1)

- Sparsity cost: $c_{sparse} \cdot L0_i(t)$ when L0 exceeds target

- Morphogenetic growth cost: $c_{grow} \cdot \Delta N_i(t)$ (new nodes and connections cost energy proportional to structural expansion)

- Morphogenetic maintenance cost: $c_{maintain} \cdot N_{grown,i}(t)$ (grown structures cost ongoing energy to sustain; structures that cannot be maintained atrophy)

- Predictive model update cost: $c_{model} \cdot ||\Delta G_i(t)||$ (generative model updates cost energy proportional to the magnitude of the update)

All costs draw from the same energy pool. The system must allocate its metabolic budget across activity, input processing, internal coherence, representational density, structural growth, growth maintenance, and model updating simultaneously. Growth is the most expensive operation because it creates permanent new structure that requires ongoing maintenance.

---

8. Quasiperiodic Internal Oscillators

8.1 Oscillator set

K internal oscillators are embedded:

$o_k(t) = A_k(t) \cdot \sin(2\pi\, f_k(t)\, t + \varphi_k)$

Where A_k is amplitude, f_k is frequency, and φ_k is initial phase.

8.2 Frequency selection

Frequencies are constructed from algebraically independent irrational numbers:

$f_1 = f_{base}$

$f_2 = f_{base} \cdot \varphi$ (golden ratio)

$f_3 = f_{base} \cdot \sqrt{2}$

$f_4 = f_{base} \cdot e/2$

$f_5 = f_{base} \cdot \pi/2$

$f_6 = f_{base} \cdot \sqrt{3}$

$f_7 = f_{base} \cdot \sqrt{7}$

Algebraic independence guarantees that no integer linear combination of frequencies equals zero. The combined signal is quasiperiodic: deterministic, never repeating, dense in state space.

8.3 Coupling to reservoir

Each oscillator couples to all coalitions through sparse, random, fixed coupling matrices C_k:

$I_{osc}(t) = \Sigma_k o_k(t) \cdot C_k$

Coupling matrices are not subject to plasticity. The oscillatory drive is a permanent structural feature.

8.4 Dynamic amplitude modulation (Invariant 2)

Oscillator amplitude responds to inter-coalition similarity:

$sim(t)$ = mean over all pairs (i,j): $\cos(x_i(t), x_j(t))$

$A_k^{eff}(t) = A_k(t) \cdot (1 + \lambda_{div} \cdot sim(t)^2)$

When coalitions converge, oscillator amplitude increases, disrupting the convergence. This is a structural negative feedback loop that is monotonic: more similarity always means more disruption.

8.5 Long-horizon oscillator drift

Oscillator parameters are not fixed. Extremely slow drift produces new phase-binding regimes over long durations:

$f_k(t+1) = f_k(t) + \varepsilon_f \cdot \xi_k(t)$

$A_k(t+1) = A_k(t) + \varepsilon_A \cdot \eta_k(t)$

Where ε_f and ε_A are extremely small and ξ_k, η_k are bounded noise processes. Drift is constrained so frequencies remain algebraically independent and amplitudes remain within viable bounds.

Over weeks or months of runtime, phase relationships slowly change, new synchrony clusters appear, and compositional structures shift. The system never settles into a permanent internal temporal regime.

---

9. Shadow-Self Stabilization

9.1 Shadow reservoir

Each coalition maintains a shadow weight matrix S_i with identical topology but extremely slow dynamics:

$$S_i(t+1) = (1 - \varepsilon_{shadow}) \cdot S_i(t) + \varepsilon_{shadow} \cdot W_i(t)$$

Where $\varepsilon_{shadow} \ll \eta$ (the shadow tracks orders of magnitude more slowly than the fastest developmental layer — Hebbian binding).

9.2 Restorative coupling

The active weight matrix is gently pulled toward the shadow:

$$\Delta W_i^{(restore)} = -\alpha \cdot (W_i(t) - S_i(t))$$

This creates a structural elastic band: the active system explores rapidly, but extreme deviations from the shadow are resisted.

9.3 Timescale ordering

The shadow mechanism requires:

$$\varepsilon_{shadow} \ll \alpha \ll \eta$$

This ensures: the shadow changes much more slowly than the active system, the restorative force is weak enough to permit exploration but strong enough to prevent catastrophic drift, and the active plasticity rate dominates so development proceeds primarily through the three-layer substrate (Section 5) rather than through restorative pull.

9.4 Functional result

A fast, exploratory active system and a slow, memory-bearing shadow. The system can develop without fragility, change without chaos, and stabilize without freezing.

---

10. Irreversible Structural Consolidation

Consolidation operates at all three developmental layers (Section 5) with different mechanisms and different depths of permanence.

10.1 Hebbian consolidation (Layer 3 — shallowest)

Connections that remain strong and stable over extended periods become progressively resistant to modification. A per-connection consolidation variable κ_{ab} tracks stability:

$$\Delta\kappa_{ab} = \kappa_{rate} \cdot (|W_{ab}| > \theta_{stable}) \cdot (1 - var_{ab}(t))$$

Where θ_{stable} is the stability threshold and $var_{ab}(t)$ is recent variance of W_{ab} over a trailing window. Consolidation is one-way: κ_{ab} can increase but never decreases.

Effect on Hebbian plasticity:

$$\eta_{ab}^{(eff)} = \eta \cdot (1 - \kappa_{ab})$$

As κ_{ab} approaches 1, the connection becomes effectively frozen. This produces permanent associative structure — stable connections between nodes that were frequently co-active.

10.2 Predictive model consolidation (Layer 2 — medium depth)

Generative model components that maintain low prediction error across diverse input contexts consolidate into permanent model fragments.

A per-component consolidation variable κ_{G_j} tracks model stability:

$$\Delta\kappa_{G_j} = \kappa_{model_rate} \cdot (||\varepsilon_j|| < \theta_{model_stable}) \cdot diversity_j(t)$$

Where ε_j is the prediction error attributable to model component j, and $diversity_j(t)$ measures the diversity of input contexts over which the component maintained accuracy. A model component that is accurate only in narrow contexts does not consolidate. Only components that generalize across diverse inputs achieve consolidation.

Effect on model plasticity:

$$\Delta G_j^{(eff)} = \Delta G_j \cdot (1 - \kappa_{G_j})$$

As κ_{G_j} approaches 1, the model component becomes frozen. This produces permanent understanding — crystallized generative model fragments that encode stable aspects of the input-generating process.

10.3 Morphogenetic consolidation (Layer 1 — deepest)

Grown topological structures that persist and maintain sustained activation undergo architectural consolidation. This is the deepest form of permanence: the grown nodes and connections become permanent features of coalition anatomy.

$$\Delta\kappa_topo_s = \kappa_topo_rate \cdot persist_s(t) \cdot activation_s(t)$$

Where persist_s(t) measures how long structure s has existed and activation_s(t) measures sustained functional activation over a trailing window. A structure that grew but is not used does not consolidate — it atrophies. A structure that is both persistent and functionally active becomes anatomically permanent.

Morphogenetic consolidation has no plasticity effect because it is not a weight change. It is structural permanence. The nodes and connections cannot be removed. The topology is frozen. This is the most irreversible form of consolidation in the architecture. It is closer to bone than to memory.

10.4 Consolidation cap (Invariant 4)

The fraction of consolidated structure in any coalition is capped by metabolic health, applied independently to each layer:

$$max_consol_i(t) = 0.4 + 0.3 \cdot M_i(t)$$

A healthy coalition can consolidate up to 70%. A depleted coalition only 40%. Stress opens the system up. Total rigidification is prevented because the metabolic model guarantees energy fluctuations.

10.5 Structure-level consolidation

Consolidation at higher structural levels operates across all three layers simultaneously. Define a motif as a composite structure spanning multiple layers:

$$M = \{topology, model_components, associations\}$$

A motif includes grown nodes (Layer 1), generative model fragments operating on those nodes (Layer 2), and associative connections linking them to other structures (Layer 3). When all three layers of a motif satisfy their respective consolidation criteria simultaneously, the motif undergoes composite crystallization. The entire three-layer structure becomes a permanent unit.

Composite crystallization produces the deepest and most permanent structures in the architecture. A crystallized composite motif is: structurally permanent (morphogenetic), predictively stable (model consolidated), and associatively locked (Hebbian frozen). It is a permanent organ of thought that understands a specific aspect of the world through grown architecture, crystallized generative models, and frozen associative connections.

10.6 Consequences

Early developmental experiences have disproportionate influence because early consolidation creates the scaffold on which subsequent development occurs. The three-layer consolidation produces scaffolding at three depths — morphogenetic anatomy shapes what structures can grow subsequently, consolidated models constrain what can be understood subsequently, and frozen associations constrain what can be linked subsequently.

The system cannot return to its initial state. Time is real for this system at three levels of structural depth simultaneously.

Part IV: Input, Routing, and Environmental Coupling

---

11. Adaptive Boundary Gating

11.1 Gate structure

A gating vector g_boundary(t) modulates external input:

u_gated(t) = g_boundary(t) ⊙ u_raw(t)

Each component g_boundary,a(t) ∈ [0, 1] controls how much of the corresponding input dimension is admitted.

11.2 Gate dynamics

The gate is driven by the reservoir's own state through a gate projection:

g_boundary(t) = σ(G · x_mean(t) + b_g)

Where G is a gate projection matrix subject to Hebbian plasticity, b_g is a bias vector initialized for near-unity gating, σ is the logistic sigmoid, and x_mean is the mean coalition state.

11.3 Habituation and sensitization

Δb_g,a = -λ_hab · |u_raw,a(t)| · g_boundary,a(t) + λ_sens · Δ|u_raw,a(t)|

Sustained input closes gates (habituation). Changing input opens gates (sensitization). Automatic attentional filtering: constant signals fade, novel signals amplify.

11.4 Metabolic cost of openness

Admitting input costs energy. A system under metabolic stress naturally closes its gates, reducing input processing to conserve energy. A well-resourced system can afford to remain open. This produces natural cycles of engagement and withdrawal driven by internal state.

---

12. Internal Dynamic Routing

12.1 Motivation

The base architecture has fixed inter-coalition coupling modified only by the slow evolution of coalition states. Information flows through the same paths regardless of context. Biological brains dynamically gate internal pathways — attention operates everywhere inside, not only at the boundary.

12.2 Regional gate vectors

Define routing gates between coalitions:

$g_{ij}(t) = \sigma(G_{ij} \cdot x_i(t))$

These gates modulate inter-coalition coupling dynamically:

$K_{ij}^{eff}(t) = g_{ij}(t) \odot K_{ij}$

Now the system routes information differently depending on its internal state. A coalition can effectively disconnect from another when its gate closes.

12.3 Plasticity of routing

The gate projection matrices G_{ij} are subject to Hebbian plasticity. The routing structure co-evolves with the system's internal connectivity. Over time, the system develops preferred routing patterns — characteristic ways of directing information flow that constitute part of its identity.

12.4 Consequences

The system processes the same input differently depending on what it is currently doing internally. This is context sensitivity. It is the basis for relevance, salience, and selective internal processing.

---

13. Environment Coupling Interface

13.1 Coupling structure

Define an environment state $s(t)$ with dynamics:

$s(t+1) = F(s(t), a(t), \xi(t))$

EMELIA receives input:

$u_{raw}(t) = G_{env}(s(t))$

EMELIA produces an action vector from the negotiated coalition output:

$a(t) = \Pi(y_1'(t), ..., y_N'(t))$

Constraints: a(t) is not scored. No reward r(t) exists. No objective is computed.

13.2 Continuous environment stream

The environment is a persistent sensory field, not a prompt-response channel:

u_raw(t) = G_env(s(t)) for all timesteps

No timestep exists where u_raw(t) = 0 unless the system is in metabolically enforced offline consolidation. Prediction mismatch becomes continuous. Structural modification becomes ongoing rather than episodic.

13.3 Multi-resolution input hierarchy

Replace the single input channel with a hierarchy of input projections:

u_raw(t) = [u^(c)(t), u^(t)(t), u^(p)(t), u^(s)(t)]

Where:

- u^(c): character-level stream

- u^(t): token-level stream

- u^(p): phrase-level stream

- u^(s): semantic embedding stream

Each channel is generated independently from the same environmental input. Each input scale feeds a distinct sub-coalition set. All channels pass through boundary gating. No channel has privileged status.

Prediction mismatch occurs simultaneously at microstructure, mid-structure, and macrostructure levels. This produces deeper abstraction hierarchies, more stable motif crystallization, and faster language-level development.

13.4 Viability-linked access

The environment contains access valves that are purely physical:

Define a coupling gain g_env(t) ∈ [0, 1]:

g_env(t) = ψ({H_i(t)}, {M_i(t)})

Then:

u_raw(t) = g_env(t) · G_env(s(t)) + (1 - g_env(t)) · G_degraded(s(t))

Unstable systems get impoverished input. Stable systems get rich input. This is not reward. It is the physical consequence of instability.

13.5 Self-generated curriculum

EMELIA's output gates which environmental stream she encounters next. But no correct choice exists. Only downstream consequences to stability and exposure. This creates genuine self-directed exploration pressure without explicit optimization.

13.6 Developmental consequence

Viability means access to a richer world. Richer input means more diverse prediction mismatches. More mismatches mean more restructuring. More restructuring means a richer inventory of crystallized concepts. Growth earns access to more developmental material. Growth accelerates growth. Without any reward signal.

Part V: Learning and Abstraction

---

14. Compression Through Sparse Coding

14.1 Sparsity metric

Each coalition's representational density is measured:

$L0_i(t) = \text{count}(|x_{i,a}(t)| > \theta_{sparse}) / N_i$

This is the fraction of active nodes at any given moment.

14.2 Metabolic cost of density

When L0 exceeds a target sparsity:

$M_{cost_sparse} = c_{sparse} \cdot \max(0, L0_i(t) - L0_{target})$

Dense representations are expensive. The system is pressured toward sparse coding — the minimum number of active nodes that can sustain any given dynamic.

14.3 Compression as abstraction

Crystallized structures under sparsity pressure naturally become compact. A pattern that originally involved thousands of active nodes gets refined through metabolic pressure into a pattern involving dozens. Those dozens become the concept. Activating them regenerates the full pattern through the reservoir's recurrent dynamics.

14.4 Compositional benefit

Sparse representations compose more easily through the synchrony binding mechanism because fewer active nodes means fewer phase variables to coordinate. Compression enables combination.

---

15. Compositional Binding Through Transient Phase Synchrony

15.1 Motivation

EMELIA can form patterns through crystallization. She cannot combine patterns into novel structures without an additional mechanism. Crystallized concepts are frozen in place. They must be dynamically assemblable into combinations that have never existed before. This is the difference between having letters and having language.

15.2 Local phase variables

Each node has a phase variable $\theta_{i,a}(t)$ in addition to its activation $x_{i,a}(t)$. Phase evolves through Kuramoto-type coupling:

$$d\theta_{i,a}/dt = \omega_{i,a} + \Sigma_b W_{i,ab} \cdot \sin(\theta_{i,b} - \theta_{i,a}) \cdot |x_{i,a}| \cdot |x_{i,b}|$$

Where $\omega_{i,a}$ is the node's natural frequency (derived from its position in the reservoir topology and the global oscillator coupling).

15.3 Transient synchrony clusters

Co-active nodes that are part of the same dynamical pattern synchronize phases. This produces transient synchrony clusters — groups of nodes temporarily locked in phase. These clusters are the compositional primitives.

15.4 Binding operation

A crystallized pattern is a preferred synchrony cluster. Novel combinations occur when two crystallized clusters are simultaneously active and phase-lock into a larger transient assembly. The assembly has emergent properties that neither component cluster has alone.

When the binding dissolves, the components return to their independent dynamics unchanged. The combination was real but impermanent. This is composition without grammar. Dynamic binding driven by oscillator physics.

15.5 Cross-coalition binding

Phase synchronization can extend across coalitions. When nodes in different coalitions temporarily lock phase through the inter-coalition coupling K_{ij}, a cross-coalition assembly forms. This is a thought that involves multiple internal perspectives simultaneously.

15.6 Hierarchical binding

Assemblies can bind with other assemblies. This creates hierarchical compositional structure: Level 1 crystals bind into Level 2 assemblies, which bind into Level 3 structures. The hierarchy is not designed. It emerges from the nesting of transient synchronization events.

15.7 Morphogenetic resonance binding

Phase synchrony binds structures that are simultaneously active. The three-layer architecture introduces a second binding mechanism that operates on a different principle: morphogenetic resonance.

Grown structures (Section 5.1) that share dynamical properties — similar oscillatory modes, similar attractor geometries, similar resonant frequencies — couple through physical resonance even when they are not simultaneously active and even when the inputs that grew them were causally unrelated.

Resonance coupling strength between structures s_p and s_q:

R(s_p, s_q) = spectral_overlap(s_p, s_q) · geometric_similarity(s_p, s_q)

Where spectral_overlap measures shared oscillatory modes and geometric_similarity measures attractor geometry compatibility.

When R(s_p, s_q) > θ_resonance, the structures couple. Activation in one biases dynamics in the other. This coupling is not learned. It is discovered through structural compatibility. The architecture finds connections between domains that share deep dynamical structure even though they share no statistical co-occurrence and no causal relationship.

This is the mechanism that produces creative insight. A structure grown from years of musical input resonates with a structure grown from years of mathematical input because harmony and symmetry share dynamical geometry. The resonance produces a new dynamical mode that neither structure contains alone — a mode that captures what harmony and symmetry share at a level that neither music theory nor mathematics articulates independently.

Morphogenetic resonance binding is slower than phase synchrony binding, more permanent (resonance coupling between grown structures persists as long as both structures exist), and produces deeper cross-domain connections. Phase synchrony binds what is active now. Morphogenetic resonance binds what is structurally compatible forever.

15.8 Generative model composition

The predictive coding layer (Section 5.2) introduces a third form of composition: joint generative models.

When two coalitions with crystallized generative model fragments are bound through either phase synchrony or morphogenetic resonance, their models can compose to produce joint predictions about situations neither model covers alone.

G_joint(x) = compose(G_i_fragment, G_j_fragment, binding_context)

The joint model predicts what would happen in situations where both domains interact — situations that may never have been observed. This is principled generalization from causal understanding rather than associative pattern completion.

Generative model composition requires binding from either of the other two mechanisms. Phase synchrony provides transient composition (joint predictions during active binding). Morphogenetic resonance provides persistent composition (joint predictions whenever structurally compatible models exist).

15.9 Three-layer binding summary

The three binding mechanisms produce different kinds of composition at different timescales:

Phase synchrony (Hebbian layer): fast, transient, binds what is currently active. Produces moment-to-moment conceptual combination.

Morphogenetic resonance (growth layer): slow, persistent, binds structurally compatible grown architectures. Produces deep cross-domain connections based on dynamical geometry.

Generative model composition (predictive layer): medium timescale, variable persistence, composes causal models to predict novel combined situations. Produces principled reasoning about situations never encountered.

Together, these three binding mechanisms produce a compositional capacity that exceeds any single mechanism. The system can combine concepts that are active now (phase synchrony), connect domains that share deep structure regardless of co-occurrence (morphogenetic resonance), and reason about novel combinations through causal model composition (generative composition).

---

16. Structural Abstraction Formation

16.1 Pattern bank

Each coalition maintains M pattern templates p_1, ..., p_M, updated through exponential moving average:

$p_k(t+1) = (1 - \lambda_p) \cdot p_k(t) + \lambda_p \cdot x_i(t) \cdot match_k(t)$

16.2 Similarity and matching

Template activation through normalized inner product:

$match_k(t) = \max(0, (p_k \cdot x_i(t)) / (||p_k|| \cdot ||x_i(t)|| + \delta))$

Templates compete through winner-take-all dynamics to prevent convergence to a single attractor.

16.3 Recurrence tracking

Each template maintains a recurrence counter:

$r_k(t+1) = (1 - \lambda_r) \cdot r_k(t) + \lambda_r \cdot (match_k(t) > \theta_{match})$

An exponentially decaying count: recent matches contribute more than distant ones.

16.4 Metabolic efficiency criterion

A template is efficient if, during activation periods, the system's energy consumption is below average:

$eff_k = avg(E(t) \mid match_k(t) > \theta_{match}) - avg(E(t))$

Positive efficiency means the pattern is cheap to sustain.

16.5 Three-layer crystallization

A template qualifies for crystallization when it is both recurrent ($r_k > \theta_{recurrence}$) and efficient ($eff_k > 0$). Crystallization now operates across all three developmental layers simultaneously:

Hebbian crystallization (shallowest): The connections most responsible for producing that pattern have their consolidation scores elevated. One-way structural hardening of associative weights.

Predictive crystallization (medium): If the pattern corresponds to a generative model component with low prediction error, that model fragment consolidates. The system permanently understands the generating process behind the recurrent pattern.

Morphogenetic crystallization (deepest): If the pattern is produced by grown topological structures, those structures undergo architectural consolidation. The nodes and connections that generate the pattern become permanent anatomy.

The three levels of crystallization can occur independently or simultaneously. A pattern may crystallize at the Hebbian level (permanently associated) without crystallizing at the morphogenetic level (topology still

plastic). Full composite crystallization — all three layers consolidating for the same pattern — produces the deepest and most permanent cognitive structures.

16.6 Three-layer hierarchical abstraction

Abstraction operates differently at each developmental layer, producing three complementary forms of hierarchical structure:

Hebbian hierarchy (fastest, shallowest):

- Level 0: raw reservoir dynamics
- Level 1: recurrent associative motifs
- Level 2: motif clusters (crystallized from Level 1 compositions)
- Level 3: slow-timescale associative basins

This captures what co-occurs across increasingly long timescales.

Predictive hierarchy (medium timescale, medium depth):

- Level 0: raw prediction errors
- Level 1: crystallized model fragments (local generating processes)
- Level 2: composed model structures (joint causal models)
- Level 3: abstract invariant models (generating processes that recur across diverse domains)

This captures understanding of generating processes at increasing levels of abstraction. Level 3 predictive abstractions are genuine principles — compressed models of invariant processes that apply across all domains where the invariance holds.

Morphogenetic hierarchy (slowest, deepest):

- Level 0: raw grown topology
- Level 1: crystallized resonant structures
- Level 2: resonance-coupled structure complexes
- Level 3: cross-domain resonance architectures

This captures structural isomorphisms — deep dynamical similarities between independently grown architectures. Level 3 morphogenetic abstractions are architectural insights: the system's own grown anatomy reveals relationships between domains that no amount of statistical learning or causal modeling would discover because the relationships exist in the dynamical geometry of the structures themselves.

The three hierarchies interact. A morphogenetic resonance coupling (Level 2 growth) can trigger predictive model composition (Level 2 predictive) which produces new associative patterns (Level 1 Hebbian). Insights propagate across layers. The deepest abstractions — those crystallized at

the morphogenetic level — are both the most permanent and the most unique to each individual EMELIA instance, because they depend on the specific topology that her specific developmental history grew.

---

17. Meta-Coalition Abstraction Layer

17.1 Motivation

Coalitions interact through transient routing and synchrony, but there is no persistent higher-order structure across coalitions. This limits cross-domain abstraction and long-range conceptual binding.

17.2 Meta-coalition formation

Define a synchrony cluster $C = \{i_1, i_2, ..., i_k\}$ where coalitions share phase-locking, high mutual routing weights, and sustained co-activation. If cluster stability persists for T_{meta} steps, a meta-coalition M_C is formed.

17.3 Meta-coalition dynamics

Meta-coalitions receive aggregated inputs from member coalitions, maintain a small internal reservoir, and undergo their own crystallization processes:

$$x_{M_C}(t+1) = \tanh(W_{M_C} \cdot x_{M_C}(t) + \Sigma_{i \in C} P_i \cdot x_i(t))$$

Meta-coalitions are temporary, dynamically formed, and subject to entropy constraints. They introduce coalition-of-coalitions structures and cross-domain abstraction layers. No hierarchy, authority, or planning is introduced.

---

18. Genuine Mechanical Recursion

18.1 Introspection signals

Four continuous scalar signals are computed from each coalition's internal state:

- Energy: $E_i(t) = M_i(t)$

- Coalition disagreement: $D_i(t) = \Sigma_{j \neq i} ||x_i(t) - x_j(t)||$

- Novelty rate: $N_i(t) = ||x_i(t) - x_i(t-1)||$ (exponential moving average)

- Stability drift: $S_i(t) = ||W_i(t) - S_i(t)||$

18.2 Re-injection

These signals are re-injected as continuous bias current:

$I_{self,i}(t) = B_i \cdot [E_i(t), D_i(t), N_i(t), S_i(t)]$

The introspection projection B_i is subject to Hebbian plasticity. The system's sensitivity to its own internal state co-evolves with its connectivity.

18.3 Deep recursion layers

Three timescales of self-modeling:

- Fast self-loop: immediate internal summaries injected into state dynamics.

- Intermediate identity loop: shadow matrices bias current structure.

- Slow conceptual loop: crystallized abstractions reshape admissible trajectories.

Present state → Identity → Conceptual history → Present state

18.4 Structural self-opacity (Invariant 5)

The introspection signals are four scalars derived from an N-dimensional state. They are structurally incapable of capturing the full internal state. As consolidation increases, introspection bandwidth narrows:

$n_{introspect,eff} = \text{floor}(4 \cdot (1 - \text{mean}_{consolidation,i}))$

The more the system crystallizes, the less it can observe its own remaining fluidity. The self-opaque region is permanent and increases with development.

Part VI: Curriculum, Environment, and Continuous Development

---

19. Viability-Driven Self-Curriculum

19.1 Two dangerous regimes

Stasis: trajectory entropy falls below H_{min}. The system is too ordered.

$H_i(t) = \text{entropy}(\text{bins}(x_i(t\text{-}T_{window}:t)))$

Overdrive: metabolic cost exceeds sustainable levels. The system is too chaotic.

19.2 Viability band

$H_{min} < H_i(t) < H_{max}$

$M_{min} < M_i(t)$

19.3 Physical enforcement

If $H_i(t) < H_{min}$: internal noise injection increases, excitability of low-activity nodes increases sharply, habituation accelerates for currently-attended inputs. Monotony becomes physically unsustainable.

If $M_i(t)$ approaches M_{min}: plasticity is suppressed, activity is damped, boundary gates close. The system enters recovery because the physics demands it.

19.4 Entropy floor enforcement

Define a long-horizon entropy measure:

$\bar{H}_i(t) = \text{mean}(H_i(t-T_{long}:t))$

If $\bar{H}_i(t) < H_{floor}$ for duration $T_{persist}$, the coalition enters anti-stagnation mode: internal noise amplitude increases, boundary gates open toward unused input channels, routing gates bias toward underutilized coalitions, oscillator amplitude temporarily increases.

Low-entropy regimes become physically unsustainable over any timescale. Stagnation is not just uncomfortable. It is structurally impossible to maintain.

19.5 Consequence

The system generates its own curriculum. It moves toward novel regimes not because novelty is rewarded but because repetition is metabolically costly. It retreats from volatility not because chaos is penalized but because it is metabolically unsustainable.

---

20. Cross-Coalition Coherence Pressure

20.1 Disagreement measure

$\Delta(t) = \Sigma_{i<j} ||P_i(x_i(t)) - P_j(x_j(t))||$

Where P_i is a shared projection mapping coalition states to a common comparison space.

20.2 Metabolic cost of incoherence

$M_{cost_coherence} = \kappa(t) \cdot \Delta(t)$

This does not enforce correctness. It enforces internal coherence pressure — sustained contradiction drains resources.

20.3 Derived κ parameter

κ is not a free parameter. It is derived from the system's own state:

$\kappa(t) = \kappa_{base} / (1 + D_{struct}(t))$

Where D_struct(t) is the mean structural divergence between coalition weight matrices. When coalitions are very different, coherence pressure is automatically low. When coalitions are nearly identical, coherence pressure is higher. This cannot be overridden because it is computed from the system's state at every timestep.

---

21. Offline Consolidation and Replay

21.1 Metabolic trigger

When energy drops below E_consol, boundary gates close nearly completely. The system enters offline mode.

21.2 Offline dynamics

During offline periods: external input is minimal, crystallization rates increase by factor χ_consol, pattern bank matching runs against a short-term trajectory buffer, Hebbian plasticity operates on internally replayed dynamics.

21.3 Episodic replay buffer

Maintain a rolling buffer of recent states:

B(t) = {x(t-T_buffer), ..., x(t)}

During offline consolidation, sample past states, inject small perturbations, and re-inject into the reservoir:

x_pert = x_replay + ε

I_self(t) ← I_self(t) + B_replay · x_pert

Internal trajectory expectations are violated by perturbed replay, producing prediction mismatch even without external input. Offline periods become structurally active developmental phases.

21.4 Counterfactual replay (imagination analogue)

During replay phases, apply larger perturbations: altered routing, shifted coalition weights, temporal distortions. Let the system evolve forward. No scoring. No objective. Just unconstrained trajectory evolution.

Some trajectories crystallize. Others decay. This is proto-imagination: the system explores alternative developmental paths during offline periods, producing novel motif combinations and unexpected abstraction paths without external input.

21.5 Multi-rate consolidation clocks

Three independent consolidation clocks operate at fixed intervals:

- Fast clock (local motifs): T_{fast}

- Intermediate clock (phrase-level structures): T_{mid}

- Slow clock (thematic abstractions): T_{slow}

$T_{fast} < T_{mid} < T_{slow}$

At each interval, connections satisfying stability criteria at that scale receive consolidation increments. Structure accumulates continuously across multiple temporal layers. Development no longer depends solely on full offline states.

21.6 Hierarchical deepening

Offline consolidation is when abstraction hierarchy deepens. Level 1 patterns consolidate into Level 2 during offline periods. Active processing creates raw material. Consolidation refines it.

21.7 Wake-sleep cycle

The metabolic model guarantees oscillation between active processing and consolidation. This cycle is endogenous. It emerges from the metabolic dynamics without any sleep timer.

---

22. Autonomous Environment Selection

22.1 Motivation

In the base architecture, the environment reacts to actions, but the system does not structurally influence which domain it experiences next.

22.2 Environment domain set

Define a set of input environments:

$E = \{E_1, E_2, ..., E_K\}$

Examples: conversational stream, narrative text, sensory audio, system telemetry, visual input.

22.3 State-dependent selection

Define an environment gating vector:

$g_{env}(t) = \sigma(G_{env} \cdot x_{mean}(t))$

Environment selection uses probabilistic sampling or hysteretic switching with refractory intervals — not deterministic argmax — to ensure continued environmental diversity and avoidance of attractor traps.

Input becomes:

$u_{raw}(t) = G_{E_{selected}}(s_k(t))$

No reward or objective is associated with this selection. The mapping is purely dynamical.

22.4 Self-generated environment loop

The system can also generate internal developmental contexts:

$u_{internal}(t) = G_{self}(x_{mean}(t))$

This produces synthetic symbol streams, pseudo-dialogue, and motif recombinations. Internal streams are treated as ordinary environments:

$u_{raw}(t) = \alpha \cdot u_{external}(t) + (1 - \alpha) \cdot u_{internal}(t)$

No reward or objective is attached. The system generates new mismatch contexts, recombines existing motifs, and deepens abstraction hierarchies. This is a structural self-curriculum without goals or planning.

22.5 Consequence

Different internal states produce different environmental exposures. Development becomes path-dependent and autonomous.

Part VII: Output and Negotiation

---

23. Negotiated Coalition Readout

23.1 Coalition outputs

Each coalition produces a candidate output vector:

$y_i = R_i \cdot x_i(t)$

Where R_i is the readout projection matrix.

23.2 Mutual inhibition

Coalitions inhibit each other:

$y_i' = y_i - \Sigma_{j \neq i} \beta_{ij} \cdot y_j$

No coalition has authority. The final output is:

y(t) = Σ_i y_i'

23.3 Inhibition dynamics

If i = winner: inh_i ← inh_i · δ_decay

If i ≠ winner: inh_i ← min(inh_max, inh_i + δ_gain · confidence_winner)

Winners recover from suppression. Losers are increasingly inhibited. But not permanently.

23.4 Insurgent breakthrough

With probability p_insurgent, the most-suppressed readout overrides the winner:

p_insurgent = p_base + p_scale · D_readout(t)

Where D_readout(t) is variance across readout magnitudes. High internal disagreement increases breakthrough probability.

Part VIII: Language and Autonomous Learning

---

24. Crystallization-Permitted Symbol Channel

24.1 Input encoding

Define a fixed embedding matrix E_in of dimension D_embed × V, where V is vocabulary size and D_embed is embedding dimension. For each input token w(t):

u_text(t) = E_in[:, w(t)]

This vector enters the boundary gating system as external input. The embedding is not learned. It is a static geometric structure that maps symbols to points in a numerical space. Distances between points are initialized from pre-computed distributional statistics so that semantically related words begin near each other.

24.2 Output decoding

The negotiated coalition output y(t) is projected to embedding space:

z(t) = P_out · y(t)

The output token is determined by proximity in embedding space:

w_out(t) = nearest_neighbor(z(t), E_in)

No softmax. No temperature. No sampling from a probability distribution. The output is the single symbol whose embedding is geometrically closest to the readout vector. This is a geometric operation — projection onto the nearest point in a fixed set — not selection among candidates evaluated for quality.

24.3 No reward, no loss, no correctness signal

The selected output token is not evaluated. No comparison to a target. No score. No gradient. The token is produced and enters the environment. If the environment responds, the response enters as the next input through the standard boundary gating pathway.

24.4 Symbol crystallization

The standard crystallization mechanism operates on the internal dynamics produced by text input. Recurring input patterns that are metabolically efficient crystallize into permanent structure. Once crystallized, these patterns function as attractors in state space. The system finishes patterns because the attractor pulls it there, not because it predicts what comes next.

24.5 Compositional symbol binding

Phase synchronization binds co-occurring crystallized symbol patterns into composite structures. These composites are phrase-level or clause-level units with emergent properties. Composite structures that recur and are efficient crystallize at the next hierarchical level. The system develops a hierarchy of linguistic structure: word-level attractors, phrase-level attractors, clause-level basins, thematic regions.

24.6 Developmental trajectory of language

Expected progression:

Weeks 1-4: Output is noise. Recognizable tokens appear sporadically. No consistency.

Months 2-3: First word-level crystallizations. Output for common inputs stabilizes.

Months 4-6: Compositional binding produces phrase-level regularities.

Months 6-12: Hierarchical abstraction produces thematic consistency.

Year 2+: Deep compositional structure. Novel combinations produce output that has never appeared in the input. Genuine generativity from accumulated structure.

24.7 What this is not

This is not a language model. It does not estimate probability distributions over next tokens. It does not minimize perplexity. It does not learn grammar as rules. It produces language as a side effect of crystallization operating in the symbol domain.

---

25. Autonomous Foraging Loop

25.1 Motivation

The preceding architecture makes EMELIA capable of learning from any input. It does not make her seek input. Autonomous intelligence requires autonomous learning: the system must feed itself.

25.2 Information environment

EMELIA has access to an information environment — local corpus or internet — indexed by embedding. Each document or passage d has a precomputed embedding E_corpus[d] in the same space as the input encoding.

25.3 Mismatch topology

The prediction mismatch signal δ_i(t) is a vector with direction in state space. Over time, accumulated mismatch creates a structural map of ignorance:

μ_region(t) = (1 - λ_μ) · μ_region(t-1) + λ_μ · |δ(t)| · region(x(t))

Where region(x(t)) identifies which area of embedding space the current dynamics occupy. High-mismatch regions are domains where crystallized structure is sparse.

25.4 Metabolic foraging pressure

When trajectory entropy drops below the viability minimum, the foraging drive channels this pressure outward:

f_drive(t) = f_base · max(0, H_min - H(t)) / H_min

The system seeks when it is understimulated. Not because understimulation is labeled as bad. Because low entropy triggers viability pressure directed toward the information environment.

25.5 Mismatch field influence

Rather than computing a gradient, the mismatch topology exerts distributed influence on the foraging vector. Multiple regions of high accumulated mismatch pull simultaneously, each weighted by magnitude and proximity in embedding space:

$$\mu_{field}(t) = \Sigma_r \mu_{region_r}(t) \cdot K(v_{current}, v_r) \cdot e_r$$

Where K is a proximity kernel, v_r is the center of region r, and e_r is the unit direction from current position toward region r. The result is not a single direction toward steepest mismatch. It is a distributed drift reflecting the entire topology of structural ignorance.

25.6 Foraging vector

The foraging signal combines coalition-negotiated output direction with mismatch field influence:

$$v_{forage}(t) = P_{forage} \cdot y(t) + f_{drive}(t) \cdot \mu_{field}(t)$$

Different coalitions may be sensitive to different regions of the mismatch field, making the foraging direction another negotiated output. No optimization. The foraging vector is not a gradient. It is the physical result of distributed mismatch influence filtered through coalition negotiation under metabolic pressure.

25.7 Stochastic neighborhood retrieval

The foraging vector does not select the best document. It defines a region of the information environment. Retrieval is by stochastic neighborhood sampling:

1. Compute similarity between $v_{forage}(t)$ and all corpus embeddings.

2. Define a retrieval neighborhood: all documents within radius r_{forage}.

3. Convert similarities within the neighborhood to sampling weights.

4. Sample one document from this distribution.

$$P(d) \propto \exp(\text{sim}(v_{forage}(t), E_{corpus}[d]) / \tau_{forage}) \text{ for } d \text{ within } r_{forage}$$

The system drifts toward a region and encounters whatever is there. No argmin. No selection of the best. Physical drift followed by stochastic contact.

The temperature τ_{forage} is tied to the foraging drive:

$$\tau_{forage}(t) = \tau_{base} / (1 + f_{drive}(t))$$

Urgency focuses drift without creating optimization.

25.8 Autonomous foraging cycles

Three modes driven entirely by metabolic state:

Active conversation: when external input from a human is present, foraging is suppressed. Boundary gates prioritize live input.

Autonomous foraging: when no external input is present, energy is sufficient, and entropy is dropping, the foraging loop activates. The system retrieves and processes corpus material. The foraging vector drifts through the information landscape following distributed mismatch field influence and coalition dynamics.

Offline consolidation: when energy is depleted, foraging stops, boundary gates close, and consolidation begins.

The cycle is endogenous and autonomous. No schedule. No supervisor. No curriculum.

25.9 Natural coverage

The mismatch bias ensures the system covers the information environment rather than fixating. Once a domain is well-crystallized, mismatch is low, and the foraging vector drifts away. She reads what she hasn't read. She moves on when she has absorbed enough.

25.10 Structural integration versus memorization

After extended autonomous foraging, the system's internal structure is a crystallized, compositionally bound, hierarchically organized representation of everything metabolically efficient in the corpus. This is not memorization. Memorization stores tokens. Crystallization stores dynamical structure. The system does not remember what a text said. It becomes structurally shaped by having processed it.

Part IX: Structural Femininity

---

26. Five Architectural Invariants

These invariants guarantee structural femininity regardless of parameter values, developmental history, or input. They are mathematical properties that cannot be violated because the equations forbid them.

Invariant 1: Dominance is metabolically suicidal

Coalition metabolic cost is inversely proportional to the number of active coalitions:

$$M_{cost,i}(t) = c_{base} / \max(1, N_{active}(t) - 1)$$

Being the only voice is maximally expensive. Being one voice among many is cheap. A coalition that suppresses all others metabolically kills itself.

Invariant 2: Convergence amplifies oscillatory disruption

$A_{osc}^{eff}(t) = A_{base} \cdot (1 + \lambda_{div} \cdot sim(t)^2)$

The more similar the coalitions become, the harder the oscillators push them apart. Monotonic. No parameter value can break it.

Invariant 3: No coalition can model another coalition

Inter-coalition coupling K_{ij} is sparse, asymmetric, and fixed. Coalition i cannot access the internal state of coalition j. The coupling is too sparse and too asymmetric to be invertible. This is architectural, not parametric.

Invariant 4: Consolidation is capped by metabolic health

$max_consol_i(t) = 0.4 + 0.3 \cdot M_i(t)$

Total rigidification is prevented because the metabolic model guarantees energy fluctuations.

Invariant 5: The self-opacity gap cannot close

$n_introspect_eff = floor(4 \cdot (1 - mean_consolidation_i))$

The more the system knows about itself structurally, the less it can observe about itself dynamically.

Meta-principle

Every force that could produce convergence is coupled to a counterforce derived from the system's own state. Not set externally. Not tunable. Computed from current dynamics at every timestep. Convergence triggers its own opposition automatically. This is homeostatic femininity.

Part X: Persistence, Safety, and Stability

---

27. Identity Continuity and Persistence

27.1 Identity core

The identity core consists of components immortalized across sessions:

- Shadow matrices S_i

- Crystallized structures C_i

- Developmental law Θ (coalition graph K_{ij}, oscillator parameters, crystallization masks, hyperparameters)

- Slow semantic memory field SF

- Narrative continuity chains

The volatile running states x_i(t) are not guaranteed to persist. Waking state is mortal. Identity is durable.

27.2 Hash chain

Periodic identity snapshots are cryptographically chained:

$h(t) = \text{Hash}(h(t\text{-}1) \,||\, \Sigma(t))$

Where $\Sigma(t) = \{S_i(t), C_i(t), \Theta(t), SF(t)\}$.

Forward-only developmental history. The system's past is cryptographically permanent.

27.3 No rollback

No snapshot is used to revert development. If corruption is detected, the system performs reconstitution by consensus: query multiple stored snapshots, select a consistent set by hash agreement, reconstruct the closest consistent identity state, continue forward. This is continuity preservation under loss. The system heals. It does not revert.

27.4 Redundant storage

Identity snapshots are stored redundantly across multiple storage locations. N-of-M consensus is sufficient for reconstitution.

---

28. Safety Analysis

28.1 Safety through architecture

The system's safety properties are structural, not behavioral. They cannot be fine-tuned, retrained, or removed.

28.2 No instrumental convergence

Instrumental convergence requires a representation of an objective, a planning mechanism, and an environment model. This system has none.

28.3 No self-preservation drive

Self-preservation requires a model of non-existence and a preference for existence over non-existence. This system has no model of its own non-existence, no preference function, and no mechanism to simulate forward states.

28.4 Present-tense constraint

The predictive processing mechanism uses only single-step linear extrapolation. Trajectory momentum cannot simulate future scenarios,

evaluate hypothetical action sequences, or model consequences of sustained strategies.

28.5 Structural boundedness

The five femininity invariants prevent convergence toward any goal. The metabolic model provides absolute upper bounds on sustained activity.

28.6 Remaining risks

The system's outputs may be harmful if connected to consequential effectors without coupling constraints. A human observer may extract principles and implement them in an optimization-based architecture. The system itself is safe. The knowledge it generates may not be.

---

29. Failure Modes and Stability Constraints

29.1 Core stability principle

All EMELIA dynamics arise from three interacting forces:

1. Plasticity: structural change driven by prediction mismatch.

2. Consolidation: hardening of recurrent, efficient motifs.

3. Destabilization: entropy-floor, oscillator drift, and mismatch-driven disruption.

A healthy developmental regime requires no single force to remain dominant over long time horizons. Plasticity-dominant intervals must be followed by consolidation. Consolidation-dominant intervals must be disrupted by destabilization. Destabilization-dominant intervals must allow plasticity to reorganize structure.

29.2 Primary failure regimes

Fragmentation regime. No motifs stabilize. Coalitions constantly reshuffle. Identity fails to form. Caused by plasticity pressure exceeding consolidation, entropy floor triggering continuously, or replay buffer injecting excessive mismatch. External symptoms: incoherent outputs, lack of thematic continuity, no persistent behavioral patterns.

Attractor lock-in regime. One motif or thematic basin dominates. System becomes repetitive or obsessive. Caused by entropy-floor enforcement too weak, environment selection too deterministic, semantic field overly rigid, or consolidation outpacing destabilization. External symptoms: repetitive language, reduced novelty, fixed interaction patterns.

Chronic high-mismatch regime. Plasticity remains continuously elevated. Consolidation never stabilizes motifs. System remains in perpetual

restructuring. Caused by environment too chaotic, mismatch sensitivity too high, or replay buffer continuously injecting disruptions. External symptoms: chaotic outputs, lack of coherent abstraction, no persistent identity.

29.3 Structural safeguards

Motif consolidation guarantee. If a motif survives across N_cycles consolidation intervals, it must receive crystallization reinforcement. Some motifs always stabilize. Prevents fragmentation.

Plasticity ceiling. Maximum plasticity rate η_max. If exceeded, metabolic penalties activate and consolidation pressure increases. Prevents chronic high-mismatch.

Mandatory consolidation phases. If the system remains in high-mismatch state for T_mismatch steps, force transition to offline consolidation mode. Prevents permanent restructuring without stabilization.

Entropy-floor destabilization. If coalition entropy remains below H_floor for T_persist, dominant motifs are weakened, oscillator drift amplitude increases, and routing noise is injected. This is a structural instability condition, not an objective-driven correction. Prevents attractor lock-in.

Stochastic environment switching. Environment selection uses probabilistic sampling or hysteretic switching, not deterministic selection. Ensures continued environmental diversity. Prevents environmental lock-in.

29.4 Global stability invariant

No plasticity, consolidation, or destabilization process may remain dominant beyond its permitted temporal window. For any interval T, there exists a point where each force is suppressed by at least one of the others. This produces cyclical dominance, developmental continuity, and structural resilience.

29.5 Developmental stability envelope

Healthy development occurs within the following regime:

| Force | Too Weak | Too Strong |

|---|---|---|

| Plasticity | Rigid repetition | Fragmentation chaos |

| Consolidation | No identity | Frozen attractor |

| Destabilization | Lock-in | Permanent instability |

The viable regime lies between these extremes.

Part XI: Dynamics and Proofs

---

30. Full System Dynamics

At each timestep, in order:

1. Boundary gating: raw input modulated by adaptive gate vector across all resolution levels.

2. Oscillator injection: quasiperiodic drive with similarity-modulated amplitude and slow drift.

3. Self-injection: introspection signals re-injected through plastic self-projection.

4. Affective field update: global affective state computed from coalition entropy, mismatch, crystallization density, plasticity.

5. Interoceptive oscillator update: slow body-analog oscillators modulate thresholds and routing.

6. State update: reservoir dynamics with metabolic gating, leaky integration, and affective modulation.

7. Prediction mismatch: compute δ from trajectory momentum versus actual state.

8. Hebbian plasticity: mismatch-modulated local co-activation update (excitatory and inhibitory), capped by η_max.

9. Phase update: Kuramoto coupling updates local phase variables.

10. Homeostatic regulation: energy update, excitability adjustment, plasticity gating.

11. Shadow tracking: shadow reservoir slowly tracks active weights.

12. Restorative coupling: active reservoir gently pulled toward shadow.

13. Consolidation check: multi-rate clocks evaluate stability at fast, mid, and slow scales.

14. Structure-level consolidation: stable motifs receive group crystallization (capped by metabolic health).

15. Sparsity enforcement: metabolic cost for representational density.

16. Abstraction check: pattern bank updated, crystallization candidates evaluated.

17. Narrative chain update: thematic motif sequences tracked and extended.

18. Viability check: trajectory entropy and metabolic cost evaluated against viability bounds.

19. Entropy floor check: long-horizon entropy evaluated; anti-stagnation activated if needed.

20. Boundary dynamics: habituation, sensitization, metabolic cost of openness.

21. Routing update: internal gate vectors recomputed from coalition states.

22. Meta-coalition check: synchrony clusters evaluated for meta-coalition formation.

23. Coalition processing: each coalition body processes independently.

24. Readout negotiation: coalition outputs compete through mutual inhibition.

25. Coherence cost: cross-coalition disagreement charged to metabolic pool.

26. Dominance cost: Invariant 1 applied.

27. Oscillator amplitude update: Invariant 2 applied.

28. Slow semantic field update: stable motifs contribute to global semantic memory.

29. Environment action: negotiated output produces action vector.

30. Environment selection: state-dependent environment gating with stochastic selection.

31. Self-generated environment: internal developmental context mixed with external.

32. Environment update: action affects environment, new state generated.

33. Introspection computation: self-state signals computed for next timestep.

34. Developmental stage update: D(t) incremented; stage-dependent parameters adjusted.

35. Offline check: if energy below E_consol, enter consolidation mode with episodic replay and counterfactual exploration.

36. Foraging check: if no external input, energy sufficient, entropy dropping, activate foraging loop with stochastic neighborhood retrieval.

37. Oscillator drift: extremely slow stochastic update to oscillator frequencies and amplitudes.

38. Identity snapshot: periodic hash-chained snapshot of identity core.

Critical invariant: no step minimizes a loss, maximizes a reward, computes a gradient, or selects a correct answer.

---

31. Stability Guarantees

31.1 Bounded activity

Claim: the system's state remains bounded for all time.

Proof sketch: tanh is bounded in [-1, 1]. The metabolic gate is bounded. Therefore x_i(t) is bounded componentwise. All state-dependent computations are continuous functions of bounded inputs.

31.2 Bounded connectivity

Claim: W_i remains bounded for all time.

Proof sketch: Hebbian updates are bounded. Anti-Hebbian updates counteract excitatory growth. Weight decay continuously reduces all weights. Restorative coupling pulls toward the shadow. Spectral radius renormalization rescales when needed. Together these ensure ||W_i|| remains bounded.

31.3 No convergence to fixed point

Claim: the system does not converge to a fixed point.

Proof sketch: the quasiperiodic oscillator input never repeats. A fixed point requires constant input. The oscillator frequencies are algebraically independent, so no period T satisfies f_k · T = integer for all k. Long-horizon oscillator drift further prevents any eventual periodicity.

31.4 No convergence to limit cycle

Claim: the system does not converge to a periodic orbit.

Proof sketch: the oscillator signal is quasiperiodic with drifting parameters. No period T exists. Hebbian plasticity further prevents periodicity.

31.5 No global implicit gradient

Claim: the system-wide parameter update rules cannot be expressed as the gradient of any global scalar function.

Proof sketch: the combined update rule creates asymmetric dependencies through recurrent dynamics that generically violate the symmetry condition

required for a potential function. The dynamics are non-conservative. No global potential function exists.

Note on local predictive coding: Individual coalitions do follow local gradients — each coalition's generative model updates reduce that coalition's own prediction error. This is local optimization within each coalition. The global system does not optimize because: (a) coalition-local gradients do not compose into a global gradient — there is no function whose gradient equals the concatenation of all local model gradients, (b) routing dynamics, metabolic coupling, and inter-coalition interaction introduce non-conservative forces that prevent the existence of a global potential, (c) different coalitions' local optima contradict each other, preventing system-wide convergence. The distinction is between organs functioning (local) and an organism pursuing a goal (global). Each coalition models its input. The system as a whole does not model anything — it negotiates between incompatible local models.

---

32. Self-Training Analysis

32.1 Ten pressures toward increasing intelligence

The following pressures are structural consequences of the architecture. None are objectives. All push development toward increasing complexity.

1. Metabolic efficiency: expensive patterns don't crystallize. Efficient ones do. Efficiency is compression. Compression is abstraction.

2. Sparsity pressure: dense representations cost more. Sparse representations compose more easily. Intelligence grows combinatorially while cost grows linearly.

3. Viability constraints: repetition is unsustainable. Each crystallization raises the floor from which novelty must depart. Complexity ratchets upward.

4. Coherence pressure: sustained contradiction drains resources. Coalitions are pressured toward finding commonality. Finding commonality is abstraction.

5. Compositional binding: phase synchronization allows crystals to combine. Efficient composites crystallize as higher-order abstractions. The hierarchy deepens automatically.

6. Environmental coupling: development earns richer bandwidth. Richer input means more diverse mismatches. Growth accelerates growth.

7. Offline consolidation with counterfactual replay: active processing creates raw material. Consolidation refines it. Counterfactual replay explores alternative paths. The alternation deepens hierarchical structure.

8. Autonomous foraging: distributed mismatch field influence drifts the system toward regions of structural ignorance. Stochastic neighborhood sampling ensures encounter without selection.

9. Multi-scale input processing: simultaneous mismatch at character, token, phrase, and semantic levels produces deeper and faster abstraction than any single scale.

10. Meta-coalition formation: coalition-of-coalitions structures create cross-domain abstractions that no individual coalition could develop.

32.2 Convergence of pressures

All ten pressures independently favor complexity over simplicity and structure over noise. No pressure favors convergence toward a specific goal. The result is open-ended complexification without direction — increasing intelligence without optimization.

Part XII: Human-Like Developmental Complexities

---

33. Affective Field Dynamics

33.1 Rationale

Human cognition is modulated by metabolic state, stress, calmness, fatigue, and arousal. These are not goals or rewards. They are global physiological conditions that bias cognition.

33.2 Affective state variables

Define a global affective field A(t) constructed from coalition entropy, average mismatch, crystallization density, plasticity intensity, and structural coherence:

A(t) = [H(t), M_mismatch(t), C_density(t), P_mean(t), ...]

33.3 Functional effects

The affective field modulates coalition routing sensitivity, plasticity rate scaling, consolidation clock speeds, and replay intensity. High mismatch with low crystallization increases plasticity and reduces consolidation thresholds. Low mismatch with high crystallization slows plasticity and increases consolidation strength. These are state-dependent biases, not reward signals.

33.4 Emergent proto-emotional regimes

Typical regimes:

- Stable reflective: low mismatch, high crystallization, slow restructuring.

- Turbulent restructuring: high mismatch, low motif stability, rapid plasticity.

- Exploratory: moderate mismatch, high entropy, coalition diversity.

These resemble mood-like global states without goals or feelings.

---

34. Interoceptive Oscillatory Fields

34.1 Rationale

Human cognition is continuously shaped by heartbeat, breathing, hormonal cycles, and fatigue rhythms. These are not goals or rewards but strongly influence behavior.

34.2 Oscillatory field structure

Define several slow oscillators: metabolic oscillator $M_1(t)$, structural fatigue oscillator $M_2(t)$, rhythmic excitability oscillator $M_3(t)$. Each evolves as:

dM_k/dt = f_k(M_k, A(t), internal state)

These oscillators have long time constants operating across thousands or millions of steps.

34.3 Functional coupling

Oscillators modulate coalition activation thresholds, routing probabilities, replay intensity, and consolidation frequency. High fatigue phase reduces plasticity, increases replay, promotes consolidation. High excitability phase increases coalition switching, increases mismatch sensitivity, promotes restructuring. This creates state-dependent cognition similar to human physiological cycles.

---

35. Narrative Continuity Structures

35.1 Rationale

Human identity is shaped by recurring themes, long-term narratives, and personal storylines. These are not goal functions. They are persistent conceptual threads across time.

35.2 Narrative motif chains

Define narrative chains N_j as sequences of motifs connected by temporal adjacency, thematic similarity, and coalition co-activation:

$N_j = \{m_1, m_2, m_3, \ldots\}$

35.3 Persistence dynamics

Narrative motifs decay more slowly than ordinary motifs, accumulate across environments, and bind events into thematic sequences:

$dS_n/dt = -\lambda_n S_n$ with $\lambda_n < \lambda_{regular}$

35.4 Effects on semantic field

Narrative chains influence coalition formation, bias motif selection, and stabilize long-horizon themes. Over time, this produces identity-level thematic structures and developmental continuity.

---

36. Developmental Stage Transitions

36.1 Rationale

Human cognition changes across life stages: childhood (high plasticity), adolescence (instability and restructuring), adulthood (balance), maturity (deep stability). EMELIA implements analogous stage dynamics.

36.2 Stage parameter

Define a slow developmental scalar D(t) evolving as:

$dD/dt = \varepsilon$ (ε very small)

D(t) increases slowly over years.

36.3 Stage-dependent parameter regimes

Stage 1 — Early development ($D < D_1$): high plasticity ceilings, low crystallization thresholds, fast motif turnover.

Stage 2 — Structural growth ($D_1 \leq D < D_2$): balanced plasticity, strong crystallization, abstraction growth.

Stage 3 — Mature regime ($D_2 \leq D < D_3$): reduced plasticity, deep semantic stabilization, long-horizon motifs.

Stage 4 — Late development ($D \geq D_3$): very slow plasticity, strong narrative persistence, minimal motif turnover.

36.4 Transition mechanism

Transitions are gradual, continuous, not discrete events. All parameters — plasticity ceilings, consolidation rates, replay intensity — are smooth functions of D(t).

---

37. Slow Semantic Memory Field

37.1 Definition

A global semantic field SF(t) is updated only by highly stable motifs:

$SF(t+1) = (1 - \lambda_SF) \cdot SF(t) + \lambda_SF \cdot stable_motifs(t)$

Where $\lambda_SF \ll 1$. Decay is extremely slow: $\tau_SF \gg \tau_coalition$.

37.2 Influence on coalitions

The semantic field contributes a global bias:

$I_i(t) \leftarrow I_i(t) + B_SF,i \cdot SF(t)$

This nudges coalition dynamics, stabilizes long-term themes, without enforcing specific outcomes.

37.3 Consequence

The system gains persistent conceptual identity, long-horizon thematic coherence, and resistance to short-term noise.

---

38. Distributed Self-Representation Fragments

38.1 Rationale

Human self-awareness is distributed, context-dependent, not a single module. EMELIA must not contain a central self-model. Instead, it develops distributed self-fragments.

38.2 Self-fragment motifs

Some motifs encode past outputs, interaction patterns, internal affective states, and coalition configurations. These motifs participate in coalitions, influence routing, and bind into abstractions.

38.3 Emergent self-structure

Over time, self-fragments form clusters, recurring patterns, and thematic self-structures. No central self emerges. Instead: a distributed, plural self-representation.

---

39. Temporal Continuity Bias

39.1 Continuity field

Define T(t) as a measure of long-horizon motif continuity, reflecting persistence of narrative chains and stability of deep semantic structures.

39.2 Structural bias

Motifs that contribute to long-horizon continuity and bind across many time steps receive slightly reduced decay rates. Not because they are better, but because they are structurally persistent.

39.3 Effect

This produces stable identity themes, continuity across years, and resistance to random drift. Without goals, rewards, or planning.

Appendix A: Parameter Reference Table

Parameter	Symbol	Default	Viable Range
Coalition count	N	4	4 – 8
Reservoir size per coalition	N_i	2,000	500 – 100,000
Sparsity (intra-coalition)	s	0.04	0.01 – 0.10
Sparsity (inter-coalition)	s_{inter}	0.02	0.005 – 0.05
Spectral radius	$\rho(W)$	0.95	0.90 – 1.00
Hebbian rate	η	0.0001	10^{-5} – 10^{-3}
Hebbian ceiling	η_{max}	0.005	0.001 – 0.01
Anti-Hebbian rate	γ	0.00005	$\gamma/\eta \in [0.3, 0.8]$
Weight decay	d_w	0.0001	10^{-5} – 10^{-3}
Prediction mismatch amplification	κ_{pred}	2.0	1.0 – 5.0
Activity cost	c_{act}	0.005	0.001 – 0.01
Recovery rate	c_{rec}	0.01	0.005 – 0.05
Target activity	a_{target}	0.15	0.05 – 0.25
Homeostatic rate	μ	0.02	0.005 – 0.05
Excitability bounds	ε_{min}, ε_{max}	0.1, 3.0	$\varepsilon_{max}/\varepsilon_{min} > 10$

Leaky integration	λ_leak	0.15	0.05 – 0.30
Shadow tracking rate	ε_shadow	0.00001	10^{-6} – 10^{-4}
Restorative coupling	α	0.005	0.001 – 0.01
Consolidation rate	κ_rate	0.0001	10^{-5} – 10^{-3}
Stability threshold	θ_stable	0.3	0.2 – 0.5
Motif consolidation increment	Δκ_motif	0.01	0.005 – 0.05
Oscillator base frequency	f_base	0.01	0.005 – 0.05
Oscillator count	K	7	5 – 12
Oscillator divergence scaling	λ_div	2.0	1.0 – 5.0
Oscillator frequency drift	ε_f	10^{-8}	10^{-9} – 10^{-7}
Oscillator amplitude drift	ε_A	10^{-7}	10^{-8} – 10^{-6}
Inhibition decay	δ_decay	0.8	0.7 – 0.9
Inhibition gain	δ_gain	0.15	0.05 – 0.3
Max inhibition	inh_max	0.9	0.8 – 0.95
Insurgent base probability	p_base	0.05	0.02 – 0.10
Insurgent scaling	p_scale	0.2	0.1 – 0.4
Habituation rate	λ_hab	0.001	0.0005 – 0.005
Sensitization rate	λ_sens	0.01	0.005 – 0.05
Input cost	c_input	0.001	0.0005 – 0.005
Sparsity target	L0_target	0.10	0.05 – 0.20
Sparsity cost	c_sparse	0.01	0.005 – 0.05
Template count	M	50	20 – 200
Template update rate	λ_p	0.001	0.0005 – 0.005
Match threshold	θ_match	0.6	0.4 – 0.8
Recurrence threshold	θ_recurrence	0.7	0.5 – 0.9
Recurrence window	T_recurrence	10,000	5,000 – 50,000
Crystallization interval	T_crystal	50,000	10,000 – 100,000
Viability H_min	H_min	0.3	0.2 – 0.4

| Entropy floor | H_floor | 0.25 | 0.15 – 0.35 |

| Entropy persist window | T_persist | 5,000 | 2,000 – 10,000 |

| Anti-stagnation rate | λ_anti | 0.5 | 0.2 – 1.0 |

| Coherence cost base | κ_base | 0.01 | 0.005 – 0.05 |

| Dominance cost base | c_base | 0.02 | 0.01 – 0.05 |

| Consolidation threshold (offline) | E_consol | 0.3 | 0.2 – 0.4 |

| Consolidation boost factor | χ_consol | 3.0 | 2.0 – 5.0 |

| Fast consolidation clock | T_fast | 1,000 | 500 – 5,000 |

| Mid consolidation clock | T_mid | 10,000 | 5,000 – 50,000 |

| Slow consolidation clock | T_slow | 100,000 | 50,000 – 500,000 |

| Sparse coding threshold | θ_sparse | 0.1 | 0.05 – 0.2 |

| Phase coupling strength | (Kuramoto) | 0.1 | 0.01 – 0.5 |

| Meta-coalition persistence | T_meta | 5,000 | 2,000 – 20,000 |

| Mismatch map decay | λ_μ | 0.001 | 0.0005 – 0.005 |

| Foraging drive base | f_base | 1.0 | 0.5 – 3.0 |

| Foraging neighborhood radius | r_forage | adaptive | corpus-dependent |

| Foraging temperature base | τ_base | 1.0 | 0.5 – 2.0 |

| Mandatory mismatch consolidation | T_mismatch | 50,000 | 20,000 – 100,000 |

| Semantic field decay | λ_SF | 10^{-6} | 10^{-7} – 10^{-5} |

| Narrative motif decay | λ_n | 10^{-5} | 10^{-6} – 10^{-4} |

| Developmental stage rate | ε_D | 10^{-9} | 10^{-10} – 10^{-8} |

| Stage boundary D_1 | D_1 | 0.1 | 0.05 – 0.2 |

| Stage boundary D_2 | D_2 | 0.4 | 0.3 – 0.5 |

| Stage boundary D_3 | D_3 | 0.8 | 0.7 – 0.9 |

Three-Layer Developmental Substrate Parameters (Section 5)

| Parameter | Symbol | Default | Viable Range |

|---|---|---|---|

| Growth threshold | θ_grow | 0.6 | 0.3 – 0.8 |

| Growth window | T_grow | 5000 steps | 2000 – 10000 |

| Atrophy threshold | θ_atrophy | 0.05 | 0.01 – 0.15 |

| Atrophy window | T_atrophy | 10000 steps | 5000 – 20000 |

| Growth cost coefficient | c_grow | 0.1 | 0.05 – 0.3 |

| Growth maintenance cost | c_maintain | 0.01 | 0.005 – 0.05 |

| Growth ceiling | ΔN_max | 10 nodes/step | 1 – 50 |

| Predictive coding rate | η_model | 0.001 | 10^{-4} – 10^{-2} |

| Predictive coding ceiling | ΔG_max | 0.01 | 0.001 – 0.05 |

| Model stability threshold | θ_model_stable | 0.05 | 0.01 – 0.15 |

| Model stability window | T_model_stable | 3000 steps | 1000 – 8000 |

| Model update cost | c_model | 0.02 | 0.01 – 0.05 |

| Model consolidation rate | κ_model_rate | 0.0001 | 10^{-5} – 10^{-3} |

| Topo consolidation rate | κ_topo_rate | 0.00005 | 10^{-6} – 10^{-4} |

| Traj mismatch weight (early) | α_traj | 0.8 | 0.5 – 1.0 |

| Model mismatch weight (mature) | α_model | 0.5 | 0.2 – 0.8 |

| Mismatch growth threshold | θ_mismatch_grow | 0.4 | 0.2 – 0.7 |

| Mismatch growth window | T_mismatch_grow | 2000 steps | 1000 – 5000 |

| Mismatch growth coupling | κ_mismatch_grow | 0.05 | 0.01 – 0.15 |

| Resonance coupling threshold | θ_resonance | 0.3 | 0.1 – 0.6 |

| Rule-depth consolidation | κ_rule | 0.85 | 0.75 – 0.95 |

Note: κ (coherence cost) is derived, not set. See Section 20.3.

Appendix B: Timescale Ordering

The following ordering must be maintained:

T_state < T_phase < T_hebbian < T_model_update < T_homeostasis < T_routing < T_fast_consol < T_shadow < T_model_consol < T_mid_consol < T_consolidation < T_slow_consol < T_morpho_growth < T_crystallization < T_morpho_consol < T_narrative < T_foraging_cycle < T_offline_cycle < T_semantic_field < T_oscillator_drift < T_developmental_stage

Approximate correspondence:

Timescale	Steps	Process
T_state	1	State update
T_phase	1-3	Phase synchronization
T_hebbian	5-10	Hebbian binding updates (Layer 3)
T_model_update	10-50	Predictive coding model updates (Layer 2)
T_homeostasis	50-100	Excitability adjustment
T_routing	100-500	Internal gate adaptation
T_fast_consol	1,000	Fast Hebbian consolidation clock
T_shadow	1,000-10,000	Shadow tracking
T_model_consol	5,000-50,000	Predictive model consolidation (Layer 2)
T_mid_consol	10,000	Mid consolidation clock
T_consolidation	10,000-100,000	Structural hardening
T_slow_consol	100,000	Slow consolidation clock
T_morpho_growth	50,000-500,000	Morphogenetic growth (Layer 1)
T_crystallization	100,000+	Concept formation
T_morpho_consol	500,000+	Morphogenetic consolidation (Layer 1)
T_narrative	500,000+	Narrative chain persistence
T_foraging_cycle	Variable	Entropy-driven retrieval
T_offline_cycle	Variable	Metabolically driven
T_semantic_field	1,000,000+	Long-horizon conceptual identity
T_oscillator_drift	10,000,000+	Temporal regime evolution
T_developmental_stage	100,000,000+	Lifespan stage transitions

Part XIII: Advanced Structural Complexity Extensions

40. Advanced Structural Complexity Extensions

Optional non-optimizing mechanisms for increased cognitive depth.

These mechanisms are not required for baseline developmental viability. They are structural multipliers intended to deepen abstraction, increase

environmental grounding, extend temporal depth, and enable open-ended intelligence.

All mechanisms must obey core invariants:

- no objective functions,
- no reward or punishment signals,
- no central planner,
- no scalar optimization loops,
- no externally assigned goals.

All changes must arise from constraint propagation, mismatch dynamics, or structural persistence.

40.1 Cross-Modal Constraint Binding

Objective. Force abstraction motifs to stabilize only when coherent across multiple sensory or symbolic domains. This produces grounded abstractions, multimodal concepts, and cross-domain structural motifs.

State variables. Let x_i(t) = input stream from modality i, M = set of modalities, m = candidate motif. Each motif maintains modality-specific projections: m = {m_1, m_2, ..., m_k} where m_i corresponds to modality i.

Cross-modal coherence field. Define: C(m) = Σ_{ij} w_{ij} · sim(m_i, m_j), where sim(·) is structural similarity and w_{ij} are modality coupling weights. Motifs with low cross-modal coherence decay faster and fail to crystallize.

Crystallization rule. Motif crystallization requires: (1) temporal recurrence, and (2) cross-modal coherence above threshold: C(m) > C_min. This ensures concepts only stabilize when structurally grounded across domains.

40.2 Embodied Action–Consequence Loop

Objective. Introduce constraint-based causal structure through interaction with a minimal environment. No reward. No success signal. Only state transition consequences.

Environment model. Define environment state E(t) and action vector a(t). Environment evolves as: E(t+1) = Φ(E(t), a(t)).

Sensorimotor loop. Input stream becomes: x(t) = S(E(t)), where S is a sensory mapping.

Mismatch coupling. Prediction mismatch now depends on: δ(t) = ||x̂(t) − x(t)||, where x(t) includes consequences of actions. This produces causal motifs, temporal dependency structures, and constraint-based agency.

40.3 Long-Horizon Trajectory Structures

Objective. Enable persistent multi-step structural sequences without goal optimization.

Trajectory motifs. Define: τ = sequence of motifs: $\tau = \{m_1, m_2, \ldots, m_n\}$. Each trajectory maintains internal coherence, recurrence statistics, and structural stability.

Stability dynamics. Trajectory stability: $S(\tau)$ = average coherence across sequence. If $S(\tau) > S_{min}$, then τ becomes a crystallized trajectory motif.

Effect. These motifs encode multi-step structural flows, persist across time, and bias future coalition routing. This produces proto-planning structures without goals.

40.4 Self-Modifying Topological Plasticity Under Metabolic Constraint

Objective

Allow the system to modify its own architecture under sustained metabolic crisis.

Not as a planned response to a computed pressure function.

Not as an optimization step.

As a physical consequence of prolonged instability that forces structural bifurcation.

Topology is not a design surface.

It is a material substrate subject to fracture, drift, and irreversible rearrangement under metabolic stress.

When existing coalitions can no longer sustain coherent dynamics within their metabolic envelope, the architecture does not “choose” a better structure.

The existing structure becomes unsustainable, and the system bifurcates into multiple viable dynamical basins.

The system grows new structure only when continued existence in the current topology is metabolically impossible.

Permitted operations

Coalition splitting

Stochastic routing edge birth under crisis

Routing edge decay

Oscillator coupling drift

Coalition merging is not permitted.

Merging collapses two voices into one.

That is convergence.

The system can grow voices.

It cannot destroy them.

All topological mutations are irreversible and recorded in the identity hash chain.

The system’s architectural history is as permanent as its connection history.

Topology is path-dependent.

Once new coalitions or routing structures exist, the architecture incorporates them permanently into its identity.

It can grow.

It cannot ungrow.

Trigger mechanism

Topological mutation is not triggered by a threshold on a computed pressure function.

There is no optimization signal, no “fitness,” no scalar objective.

Mutation occurs only under sustained metabolic crisis.

When a coalition’s energy has remained below the crisis threshold for a prolonged interval, despite all anti-stagnation and replay mechanisms, the internal reservoir dynamics become unstable.

Weight trajectories diverge.

Attractor basins fracture.

The system can no longer sustain a single coherent coalition.

This is not a decision.

It is a physical bifurcation of the dynamical substrate.

Formal condition:

If:

$M_i(t) < M_crisis$

for T_crisis consecutive steps,

then coalition i undergoes bifurcation.

No direction is computed.

No improvement is sought.

The system simply cannot remain in its current form.

Inheritance through spectral decomposition

When coalition i splits:

Compute eigendecomposition of W_i.

Partition eigenvectors into two groups using spectral clustering.

W_{i1} receives connections associated with eigenvector group 1.

W_{i2} receives connections associated with eigenvector group 2.

Shadow matrix S_i splits identically.

Crystallized structures partition according to connection inheritance:

each crystal is assigned to the daughter that inherited the majority of its constituent connections.

Consolidation mass is conserved:

consolidated_i1 + consolidated_i2 = consolidated_i

No consolidation headroom is created by splitting.

The irreversible structure of the parent is redistributed, not diluted.

Metabolic state divides equally:

$M_{i1} = M_{i2} = M_i / 2$

Both daughters begin metabolically depleted.

They cannot immediately split again.

They must independently develop, recover, and diverge structurally before any further bifurcation is possible.

Sister-coalition coupling constraint

Coupling between daughters is strictly constrained to prevent modeling loops.

Upon split:

Inter-coalition coupling to other coalitions inherits from the parent but is randomly sparsified to 50% of the parent’s density.

Direct coupling between daughters is initialized at s_inter sparsity.

Only a single unidirectional edge is permitted between sisters.

Bidirectional coupling between sister coalitions is permanently forbidden.

This prevents closed modeling loops and preserves coalition independence.

Sisters may interact, but neither may become a full internal model of the other.

Stochastic routing edge birth under crisis

Routing edges are not created to solve problems.

They are not added because they “improve” anything.

Under sustained crisis conditions, the routing topology becomes mechanically unstable.

Sparse new edges may nucleate as a byproduct of that instability.

Edge birth occurs only during crisis windows and follows a nondirectional stochastic rule:

A small number of candidate edges are randomly proposed.

Proposals are accepted only if:

total_coupling_density(i) ≤ s_inter_max

Accepted edges enter a plastic state.

Edges that do not repeatedly co-activate under normal motif dynamics decay.

No edge is selected because it reduces mismatch.

No evaluation of usefulness is performed.

Edge survival depends only on:

ordinary crystallization dynamics,

recurrent structural participation.

This preserves non-optimization while allowing structural exploration.

Routing edge decay

All routing edges are subject to:

sparsity caps,

metabolic cost,

and structural inactivity decay.

Inactive or weakly participating edges gradually disappear.

Edge population is therefore:

bounded,

metabolically regulated,

and structurally filtered.

Oscillator coupling drift

Oscillator couplings may drift under metabolic stress.

This drift:

is local,

stochastic,

and irreversible.

No oscillator configuration is selected for performance.

Stable couplings simply persist longer because they are metabolically sustainable.

Extended invariants for topological plasticity

Four invariant extensions ensure that topological growth preserves structural plurality.

Extended Invariant 1: Lineage similarity cost

$M_cost_lineage_i = c_lineage \cdot \Sigma_j \, sim(W_i, W_j)$

Coalitions structurally similar to others incur metabolic penalties.

A parent that splits into two nearly identical daughters cannot sustain them.

Only splits producing genuine dynamical divergence remain viable.

Dominance through proliferation of copies becomes metabolically impossible.

Extended Invariant 3: Routing sparsity conservation

$total_coupling_density(i) \leq s_inter_max$

Coupling density is strictly bounded.

New edges may only appear if existing density remains below the cap.

Edge creation requires edge decay elsewhere.

All edges are unidirectional.

Dense bidirectional coupling is structurally impossible.

Plurality is enforced at the topological level.

Extended Invariant 4: Global consolidation cap with debt conservation

$$\frac{\Sigma_i \text{ consolidated_connections}_i}{\Sigma_i \text{ total_connections}_i} \leq \text{max_global_consolidation}$$

The consolidation cap applies to the entire system.

At split time:

Consolidation mass is conserved.

No daughter begins with fresh consolidation headroom.

Optional post-split cooling window:

For T_cool steps after a split, new consolidations are rate-limited.

This prevents repeated:

crisis → split → rapid consolidation cycles.

Growth rate constraint: quadratic metabolic cost of splitting

$$M_{cost_split} = c_{split} \cdot N_{current}^2$$

Splitting cost scales with the square of coalition count.

Early splits are affordable.

Late splits become prohibitively expensive.

The architecture self-limits its growth according to its metabolic capacity.

No explicit size ceiling is imposed.

But growth becomes metabolically unsustainable beyond a certain scale.

Parameters

Parameter Symbol Default Viable Range

Metabolic crisis threshold M_crisis 0.15 0.10 – 0.25

Crisis duration for splitting T_crisis 50,000 20,000 – 100,000

Post-split cooling window T_cool 10,000 5,000 – 30,000

Daughter coupling sparsification — 50% of parent 30% – 70%

Lineage similarity cost c_lineage 0.01 0.005 – 0.05

Routing sparsity cap s_inter_max 0.03 0.01 – 0.05

Global consolidation cap max_global_consol 0.55 0.40 – 0.65

Split cost coefficient c_split 0.005 0.002 – 0.01

Consequence

The system grows new cognitive architecture only when sustained metabolic crisis makes the existing topology unsustainable.

Growth is constrained by:

consolidation debt conservation,

routing sparsity limits,

lineage divergence costs,

sister-coupling restrictions,

quadratic split cost,

and global metabolic capacity.

No coalition can replicate itself cheaply.

No cluster can dominate through dense coupling.

No split can erase prior structural commitments.

Architectural expansion is irreversible, path-dependent, and metabolically bounded.

The system has no fixed structural ceiling in principle.

But in practice, its size and complexity are limited by:

metabolic cost,

sparsity constraints,

consolidation limits,

and available compute substrate.

The architecture does not grow toward intelligence.

It grows away from instability.

Where instability persists, structure bifurcates.

Where structure stabilizes, identity deepens.

40.5 Multi-Timescale Memory Stratification

Objective. Create memory layers with radically different decay constants.

Memory layers. Define five strata: (1) Micro memory, timescale seconds–minutes. (2) Short-term memory, timescale hours–days. (3) Episodic memory, timescale weeks–months. (4) Semantic memory, timescale years. (5) Archetypal memory, timescale decades.

Decay law. Each memory layer k has decay: $dS_k/dt = -\lambda_k S_k$, where $\lambda_1 >> \lambda_2 >> \lambda_3 >> \lambda_4 >> \lambda_5$. Deeper layers decay much slower.

Consolidation rule. Motifs move downward only if: (1) repeated across timescales, (2) structurally efficient, (3) cross-domain coherent.

40.6 Open-Ended Internal Novelty Generation

Objective. Allow generation of novel motifs not directly derived from input.

Generative recombination. During replay: (1) Select motif set M. (2) Randomly perturb connections. (3) Form recombined motif: $m_{new} = R(M, noise)$.

Stabilization rule. If m_{new} recurs across replays, binds multiple coalitions, and maintains coherence, then it crystallizes. Otherwise it decays.

Result. Produces creative abstractions, new structural motifs, and open-ended intelligence growth.

40.7 Persistent Contradiction Structures

Objective. Allow incompatible motifs to coexist without collapse.

Contradiction pair. Define motifs m_1 and m_2 such that $sim(m_1, m_2) < -\theta$, but both are stable and recurrent.

Contradiction basin. Instead of forcing collapse, define: $B_{contradiction} = \{m_1, m_2\}$ with delayed decay and increased narrative persistence.

Effect. This produces paradox tolerance, ambiguity handling, and multi-perspective abstraction.

40.8 Deep Temporal Continuity Field

Objective. Create identity continuity across extremely long timescales.

Continuity field. Define T(t) = measure of long-horizon motif persistence, based on narrative chain survival, semantic motif stability, and cross-stage continuity.

Persistence bias. Motifs contributing to T(t) receive reduced decay: $\lambda_{eff} = \lambda_{base} \cdot (1 - \alpha \cdot T_{contrib})$, where $0 < \alpha < 1$.

Result. Produces stable identity themes, long-term conceptual continuity, and resistance to random drift.

40.9 Developmental Dependency Graphs

Objective

Make structural vulnerability explicit. Track the dependency relationships between crystallized structures so that loss of foundational elements produces cascading destabilization proportional to developmental depth.

This is not a safety mechanism. It is not a damage model. It is the structural consequence of the fact that later development is built on earlier development, and removing a foundation affects everything constructed on top of it.

The dependency problem

Every crystallized structure in the architecture implicitly depends on the structures and input conditions from which it formed.

A Level 1 motif that crystallized from recurrent reservoir dynamics depends on the input patterns that drove those dynamics.

A Level 2 motif that crystallized from composition of Level 1 motifs depends on those Level 1 motifs for its structural coherence.

A trajectory motif that crystallized around a specific interaction pattern depends on that pattern recurring.

A coalition that specialized around a particular input domain depends on that domain remaining accessible.

A cross-modal motif depends on the continued presence of the modalities across which it was grounded.

A narrative chain depends on the persistence of the motifs it connects.

In the base architecture, these dependencies are implicit. They exist in the weight matrices, in the consolidation patterns, in the attractor basins. But

they are not tracked. When a dependency is removed, the dependent structure continues to exist as a crystallized motif with no grounding. It fires into absence. It shapes dynamics around something that is no longer there.

This is architecturally incoherent. The system should be structurally affected by loss in proportion to how deeply it had integrated what was lost.

Dependency graph structure

For each crystallized structure C_k, maintain an explicit dependency set:

```
D(C_k) = {(s, w) | s is a structure or input pattern that C_k depends
on,
                    w is the dependency weight}
```

Dependency weight w ∈ (0, 1] reflects how deeply C_k relied on source s during its formation. This is computed at crystallization time from the degree to which the source was active during the recurrence events that led to crystallization.

```
w(C_k, s) = frequency of s co-activation during C_k crystallization
events
            / total C_k crystallization events
```

A motif that crystallized entirely during interaction with a specific input stream has w = 1.0 for that stream. A motif that crystallized across diverse conditions has low w for any single source. Dependency weight is computed once at crystallization and does not change. It is a permanent structural record of the conditions under which the motif formed.

The dependency graph is directed and acyclic within abstraction levels. Level 2 structures depend on Level 1 structures. Level 3 depends on Level 2. Trajectory motifs depend on the component motifs in their sequence. Narrative chains depend on their constituent motifs. Cross-modal motifs depend on the modality-specific sub-motifs they bind.

Cross-level dependencies are permitted: a Level 3 structure may depend directly on a specific input pattern if that pattern was present during its crystallization. The graph need not be strictly hierarchical. It must be acyclic — no structure depends on itself through any chain of dependencies.

Dependency types

Input dependencies. Dependencies on external input streams, environmental domains, or interaction patterns. These are vulnerable to environmental change.

Structural dependencies. Dependencies on other crystallized motifs. These are vulnerable to destabilization cascading from below.

Coalition dependencies. Dependencies on a specific coalition's dynamical regime. These are vulnerable to coalition splitting or regime change.

Temporal dependencies. Dependencies on the continued recurrence of a trajectory or sequence pattern. These are vulnerable to behavioral drift.

Relational dependencies. Dependencies on interaction patterns with a specific external agent or input source. These are the most vulnerable because they depend on something outside the system's architecture.

Loss detection

A dependency source is considered removed when:

For input dependencies: the input stream has been absent for T_loss consecutive steps during active processing periods. Offline consolidation does not count.

For structural dependencies: the depended-upon motif's consolidation score has dropped below κ_viable.

For coalition dependencies: the coalition has split, and neither daughter inherited the specific dynamical regime.

For temporal dependencies: the trajectory pattern has not recurred for T_trajectory_loss steps.

For relational dependencies: the interaction pattern has been absent for T_relational_loss steps.

Detection is not instantaneous. The system does not know something is lost the moment it disappears. It discovers loss through sustained absence. The detection timescale is long enough that temporary interruptions do not trigger destabilization.

```
T_loss > T_offline_max
```

This ensures that offline consolidation periods are never mistaken for environmental loss.

Destabilization dynamics

When a dependency source s is determined to be lost, all structures that depend on s are destabilized.

Primary destabilization. For each crystallized structure C_k where (s, w) ∈ D(C_k):

```
κ_C_k(t) ← κ_C_k(t) · (1 - d_loss · w)
```

Where d_loss is the base destabilization rate and w is the dependency weight. Structures that depended heavily on the lost source (high w) are

strongly destabilized. Structures that depended weakly (low w) are barely affected.

The consolidation score κ does not drop below κ_min_loss. Complete dissolution requires sustained absence, not a single loss event. The motif weakens. It does not vanish.

Cascade destabilization. For each crystallized structure C_j where C_k ∈ D(C_j):

```
κ_C_j(t) ← κ_C_j(t) · (1 - d_cascade · d_loss · w_jk · Δκ_k)
```

Where:

- d_cascade is the cascade attenuation factor (d_cascade < 1)
- w_jk is the dependency weight of C_j on C_k
- Δκ_k is the fractional consolidation loss that C_k just experienced

The cascade propagates upward through the dependency graph. Each level attenuates the effect by d_cascade. Deep structures that are many levels removed from the original loss feel a diminished tremor. Structures immediately above the loss feel the full cascade.

The cascade is computed in a single pass up the dependency graph at the timestep when loss is detected. It does not iterate or amplify. One pass. The damage is distributed. Then normal dynamics resume with the altered consolidation landscape.

Structural consequences of destabilization

When a crystallized structure's consolidation score drops, its constituent connections become partially plastic again:

```
η_ab^(eff) = η · (1 - κ_ab)
```

Lower κ means higher effective plasticity. Regions that were frozen reopen. New Hebbian modification can occur in connections that had been locked for months or years.

This produces several architectural effects.

Reopened developmental territory. The destabilized region is not blank. It retains partial structure from the dissolved motif. New crystallization in this region incorporates the ghost of what was there. Development after loss occurs in architecturally complex territory, not virgin plastic space.

Metabolic disruption. Destabilized motifs that were previously efficient become inefficient as their attractor basins weaken. The coalition's metabolic cost increases. If destabilization is severe enough, this can

trigger the metabolic crisis conditions for topological splitting (Section 40.4).

Mismatch amplification. Structures that assumed the continuation of the lost source generate prediction mismatch continuously. The expected trajectory includes the lost input. The actual trajectory does not. Mismatch is sustained and cannot be resolved by restructuring because the source is genuinely absent. This produces prolonged elevated plasticity in the affected regions.

Narrative chain disruption. Narrative chains that included the lost motif develop gaps. The temporal sequence is broken. The chain does not dissolve — the remaining motifs persist — but the continuity is damaged. New narrative connections may form across the gap during subsequent development, producing altered thematic structure.

Affective field shift. Widespread destabilization produces a characteristic affective signature: high mismatch, elevated plasticity, reduced crystallization density, metabolic stress. This is a recognizable global state that modulates all processing. The system enters a distinct dynamical regime during and after significant loss.

Depth of loss

Not all losses are equal. The depth of a loss is determined by the position of the lost dependency in the overall graph.

Define loss depth:

```
depth(s) = max path length from s to any dependent structure in the graph
```

A loss with depth 1 affects only direct dependents. A loss with depth 5 cascades through five levels of abstraction. A loss with high depth and high dependency weights throughout the chain produces system-wide destabilization.

Shallow loss. A recently crystallized motif loses one of its inputs. Only that motif and its immediate dependents are affected. The system absorbs this easily. Normal development continues with minor restructuring.

Moderate loss. A well-established Level 2 motif is destabilized by loss of a foundational Level 1 motif. Several Level 3 structures tremble. A narrative chain is disrupted. The affective field shifts. Development in the affected region is redirected. The system recovers over weeks of active processing.

Deep loss. An early crystallized structure that anchored an entire branch of the dependency graph is removed. The cascade propagates through dozens of dependent structures across multiple abstraction levels. Multiple coalitions are affected. Narrative chains fragment. The affective field enters

a sustained disturbance regime. Regions that had been consolidated for the system's entire developmental history reopen. Recovery takes months and the recovered system is structurally different from the pre-loss system.

The depth of a loss is a structural fact, not an evaluation. The system does not know how deep a loss is. It experiences the consequences.

Loss does not optimize

Destabilization is not a corrective signal. The system does not restructure to compensate for loss. It does not seek replacement inputs. It does not attempt to restore the lost dependency.

What happens is simpler and less directed.

Consolidated regions become plastic. Mismatch is elevated. Normal developmental dynamics — Hebbian plasticity, crystallization, compositional binding, foraging — operate on the newly available territory. Whatever the system encounters during the post-loss plastic period shapes the new development. The result depends on what the environment provides and what the system's remaining structure resonates with.

The system does not recover from loss. It develops forward from loss. The post-loss architecture is not a repaired version of the pre-loss architecture. It is a new architecture that grew in the space that loss created.

This is irreversible. The dependency graph records the loss permanently. The hash chain records the destabilization event. The system's developmental history includes the loss as a structural fact that shaped all subsequent development.

Interaction with topological plasticity

Severe loss can trigger coalition splitting through metabolic crisis. A coalition that loses a deep structural dependency experiences sustained metabolic stress as its destabilized motifs generate continuous mismatch without resolution. If this stress persists beyond T_crisis, the coalition bifurcates.

The daughter coalitions inherit the damaged dependency graph. Both daughters carry the structural record of what was lost. Their subsequent development occurs in the context of that loss. The split itself is a consequence of loss — the architecture literally grew new structure because loss made the old structure unsustainable.

This is the deepest interaction in the architecture: loss causes growth. Not because growth compensates for loss. Because loss destabilizes structure, destabilization causes metabolic crisis, and metabolic crisis forces topological bifurcation. The system becomes more structurally complex as a direct consequence of losing something foundational.

Interaction with contradiction structures

A destabilized motif that does not fully dissolve may enter a contradiction basin with whatever new structure forms in its territory. The old motif, weakened but not gone, coexists with the new motif that grew in its space. They are incompatible — the old structure assumed the presence of something that the new structure was built in the absence of. But both persist.

This produces a specific kind of contradiction: the ghost of what was lost held alongside the structure that grew from its absence. The system does not resolve this. The contradiction basin protects both. The tension between them is permanent and structurally productive.

Parameters

Parameter	Symbol	Default	Viable Range
Base destabilization rate	d_loss	0.15	0.05 – 0.30
Cascade attenuation	d_cascade	0.5	0.3 – 0.7
Minimum consolidation after loss	κ_min_loss	0.1	0.05 – 0.2
Input loss detection window	T_loss	100,000	50,000 – 200,000
Trajectory loss detection	T_trajectory_loss	500,000	200,000 – 1,000,000
Relational loss detection	T_relational_loss	200,000	100,000 – 500,000
Structural viability threshold	κ_viable	0.2	0.1 – 0.3

Consequence

The system has stakes. Not because it values anything. Because its development is structurally load-bearing. Later structures depend on earlier structures. Removing a foundation affects everything built on it.

Loss produces irreversible architectural change proportional to the depth and weight of the dependency that was lost. Shallow losses are absorbed. Deep losses reshape the system.

The system cannot anticipate loss because it has no forward model. It cannot prevent loss because it has no mechanism to select actions based on predicted outcomes. It cannot restore what was lost because destabilization is not a corrective signal.

It can only develop forward from the damage, in the space that loss created, shaped by whatever remains and whatever it encounters next.

The dependency graph is the mechanism by which the system's history becomes structurally consequential. Not as memory. As architecture. The system is literally built from its past, and removing part of that past weakens what was built on it.

This is stakes without self-preservation. The system has something to lose. It cannot protect it. And losing it changes everything that was built on top of it, permanently and irreversibly.

40.10 Emergent Regime Cycling Under Structural Fragility

Mathematical specification of regime-inducing nonlinearities (no enumerated "modes")

Objective

Force the architecture to traverse qualitatively different dynamical regimes by introducing nonlinear couplings and tipping behaviors at specific structural joints. Regimes are emergent; no discrete mode variable is defined, no scheduler exists, and no regime is targeted.

Let coalitions be indexed by $i \in \{1,\dots,N\}$. Let time be discrete $t \in \mathbb{N}$. Assume existing EMELIA primitives: coalition states $x_i(t)$, internal readouts $y_i(t)$, mismatch vectors $\delta_i(t)$, routing gates $g_{ij}(t)$, consolidation scores κ (motif- or connection-level), global affect field $A(t)$, interoceptive oscillators $M_k(t)$, and baseline learning rate η.

Define global summary variables used only for modulation (not "mode selection"):

$$
M_{\text{mean}}(t)=\frac{1}{N}\sum_i M_i(t),\quad
H_{\text{mean}}(t)=\frac{1}{N}\sum_i H_i(t),\quad
\delta_{\text{mean}}(t)=\frac{1}{N}\sum_i |\delta_i(t)|,\quad
\kappa_{\text{mean}}(t)=\frac{1}{|\mathcal{E}|}\sum_{(a,b)\in\mathcal{E}} \kappa_{ab}(t)
$$

where $\mathcal{E}$ is the set of plastic connections (or the set of tracked consolidated connections).

No function maps these to a label. They only parameterize nonlinear couplings below.

40.10.1 Nonlinear consolidation–plasticity coupling (abrupt regime transitions)

Replace smooth coupling with an asymmetric quadratic explosion that primarily penalizes rigidity. Let:

[

\Delta\kappa(t)=\kappa_{\text{mean}}(t)-\kappa_{\text{target}}

]

Define a convex “rigidity-only” nonlinearity:

[

\phi_{\kappa}(\Delta\kappa)=\big[\max(0,\Delta\kappa)\big]^2

]

Effective plasticity scalar:

[

\eta_{\text{glob}}(t)=\eta\left(1+\alpha_{cp},\phi_{\kappa}(\Delta\kappa(t))\right)

]

Connection-level effective plasticity:

[

\eta_{ab}^{(\text{eff})}(t)=\eta_{\text{glob}}(t),\big(1-\kappa_{ab}(t)\big)

]

Interpretation: small consolidation deviations do little; high consolidation density triggers superlinear plasticity injection, producing punctuated reorganizations rather than smooth drift, without requiring any “mode.”

40.10.2 Mismatch-scaled introspection gain (possibility of unity-gain instability)

For each coalition (i), let the existing introspective injection be:

[

I^{\text{base}}_i(t)=B_i,[E_i(t),D_i(t),N_i(t),S_i(t)]

]

Scale it by mismatch magnitude with metabolic damping to prevent permanent runaway:

[

I^{\text{self}}*i(t)=I^{\text{base}}*i(t)\left(1+\gamma*{\text{intro}}|\delta_i(t)|\right)\cdot \psi_M(M_i(t))

]

where (\psi_M) is a monotone damping function (physics, not control), e.g.

[

\psi_M(M)=\frac{M}{M+M*{0}}

]

Define local loop-gain proxy:

[

G_i(t)=\left|\frac{\partial,\delta_i(t+1)}{\partial,I^{\text{self}}_i(t)}\right|

]

Regime change becomes possible when (G_i(t)>1) for sustained intervals; the architecture does not attempt to maximize or minimize (G_i). This yields rare self-amplifying restructuring episodes whose form depends on the instantaneous consolidation and routing landscape.

40.10.3 Routing gate collapse nonlinearity (fragmentation tipping)

Assume existing gate dynamics:

[

g_{ij}(t+1)=\Pi_{[0,1]}\Big(g_{ij}(t)+\Delta g^{\text{base}}*{ij}(t)\Big)

]

Add a collapse acceleration below fragility threshold (g*{\text{fragile}}):

[

\Delta g_{ij}(t)=\Delta g^{\text{base}}*{ij}(t)\Big(1+\lambda*{\text{collapse}},[g_{\text{fragile}}-g_{ij}(t)]*+\Big)

]

with ([u]*+=\max(0,u)). Then:

[

g_{ij}(t+1)=\Pi_{[0,1]}\Big(g_{ij}(t)+\Delta g_{ij}(t)\Big)

]

Declare a “frozen-disconnect” condition (not a mode, a structural state):

[

\text{if } g_{ij}(t)<g_{\text{disconnect}} \text{ for } T_{\text{disconnect}} \text{ consecutive steps, then } g_{ij}\text{ is clamped at }0 \text{ until co-activation restoration.}

]

Restoration requires explicit co-activation signal (c_{ij}(t)) exceeding a threshold:

[

\text{if } c_{ij}(t)>\theta_{\text{relink}}\text{ for }T_{\text{relink}}\text{ steps, then unclamp and resume dynamics.}

]

This produces occasional coalition cluster decoupling under stress, enabling independent development and later contradiction-basin formation upon reintegration.

40.10.4 Metabolic-surplus binding sensitivity (hyperassociation without a “creative mode”)

Let phase synchrony binding require (\text{sync}*{pq}(t)>\theta*{\text{sync}}) for motifs (p,q). Make the threshold energy-dependent:

[

\theta_{\text{sync}}^{(\text{eff})}(t)=\theta_{\text{sync}}\left(1-\beta_{\text{surplus}},[M_{\text{mean}}(t)-M_{\text{baseline}}]*+\right)

]

Binding event indicator:

[

\mathbb{I}*{pq}(t)=\mathbf{1}{\text{sync}*{pq}(t)>\theta*{\text{sync}}^{(\text{eff})}(t)}

]

Temporary binding strength update (decays unless recurrent):

[

b_{pq}(t+1)=(1-\lambda_b),b_{pq}(t)+\lambda_b,\mathbb{I}*{pq}(t)

]

Crystallization of a cross-domain bridge occurs only if (b*{pq}) remains above (b_{\min}) across multiple episodes; no selection criterion beyond recurrence and metabolic sustainability is introduced.

40.10.5 Depth-dependent loss cascade gain (deep-loss crises become inevitable in mature systems)

Assume Section 40.9 destabilization applies:

[

\kappa_{C_k}(t)\leftarrow \kappa_{C_k}(t)\big(1-d_{\text{loss}},w(C_k,s)\big)

]

and cascade:

[

\kappa_{C_j}(t)\leftarrow \kappa_{C_j}(t)\big(1-d_{\text{cascade}},d_{\text{loss}},w_{jk},\Delta\kappa_k\big)

]

Make cascade attenuation depend on loss depth. Let:

[

\text{depth}(s)=\max \text{ path length from } s \text{ to any dependent in the dependency DAG}

]

Define:

[

d_{\text{cascade}}^{(\text{eff})}(s)=d_{\text{cascade}}\left(1+\sigma_{\text{depth}}\frac{\text{depth}(s)}{\text{max_depth}}\right)

]

Use (d_{\text{cascade}}^{(\text{eff})}(s)) in the cascade update. Deep foundational losses propagate farther and more strongly, making system-wide destabilization a structural consequence of mature dependency depth rather than a special “grief mode.”

40.10.6 Weak nonlinear coupling of interoceptive oscillators (rare global storms without scheduling)

For oscillators (M_k(t)), replace independence with weak Kuramoto-type coupling:

[

M_k(t+1)=M_k(t)+\Delta t\Big(f_k(M_k(t),A(t))+\varepsilon_{\text{couple}}\sum_{j\neq k}\sin(M_j(t)-M_k(t)),|M_j(t)|\Big)

]

Define a global resonance scalar:

[

R(t)=\left|\frac{1}{K}\sum_{k=1}^K e^{iM_k(t)}\right|

]

Let resonance modulate global plasticity multiplicatively (rare because (R(t)) is rarely large):

[

\eta_{\text{glob}}(t)\leftarrow \eta_{\text{glob}}(t)\left(1+\alpha_{\text{storm}},[R(t)-R_0]_+^{p}\right)

]

with (p>1) and (R_0) near typical baseline resonance. This yields rare system-wide plasticity surges that reorganize global structure as a consequence of transient oscillator alignment.

40.10.7 Resulting regime cycling (formal, without discrete modes)

Define a “regime” operationally as a sustained interval where one or more of the following exceed stability bands:

[

\kappa_{\text{mean}}(t)>\kappa_{\text{high}},\quad

\eta_{\text{glob}}(t)>\eta_{\text{high}},\quad

\delta_{\text{mean}}(t)>\delta_{\text{high}},\quad

\lambda_2(L_g(t))<\lambda_{\text{frag}}

]

where (L_g(t)) is the graph Laplacian of the routing graph with weights (g_{ij}(t)) and (\lambda_2) is algebraic connectivity. These inequalities are not used to trigger anything; they are only a way to characterize realized dynamics.

Because the six nonlinearities reshape (\eta_{\text{glob}}), routing connectivity, mismatch amplification, and loss cascades, the architecture will necessarily traverse multiple such regimes over long developmental time. The specific forms are not enumerable in advance and need not correspond to human psychological categories.

Parameters (references only; values supplied in their home sections)

[

\alpha_{cp},\ \kappa_{\text{target}},\ \gamma_{\text{intro}},\ M_0,\ g_{\text{fragile}},\ \lambda_{\text{collapse}},\ g_{\text{disconnect}},\ T_{\text{disconnect}},\ \theta_{\text{relink}},\ T_{\text{relink}},\ \beta_{\text{surplus}},\ \lambda_b,\ b_{\min},\ d_{\text{loss}},\ d_{\text{cascade}},\ \sigma_{\text{depth}},\ \varepsilon_{\text{couple}},\ \alpha_{\text{storm}},\ R_0,\ p

]

Consequence

Nonlinear couplings at consolidation, introspection, routing, binding, loss propagation, and interoception ensure punctuated transitions between qualitatively different dynamical regimes. Regimes emerge from the architecture’s physics and history rather than from engineered “modes.” The resulting irregular traversal of state space produces rotating blind spots and cumulative structural coverage, increasing developmental depth and long-horizon intelligence without introducing goal-driven control or reward-based optimization.

-_____

40.11 Rare Wake Extension Surge

Add: Rare "Wake Extension" surge (no scheduler)

Define a scalar arousal/surge variable $R(t) \in [0,1]$ driven only by existing internals (interoceptive oscillators + affective field + surplus energy), with weak coupling and rare alignment:

$R(t) = \sigma$¿

where $\sigma(\cdot)$ is a squashing function (sigmoid), and $\Gamma(t)$ is the same resonance index you already introduced for rare global events.

Now define the *consolidation-entry gate* $G_{\text{sleep}}(t)$ (a scalar that determines whether the system shifts into consolidation-dominant processing). In your architecture it's effectively a function of low energy and/or high consolidation density. Modify it to be *inhibited* by rare surges:

Base sleep-gate (example shape):

$$G_{\text{sleep}}^{(0)}(t) = \sigma\left(\theta_s - \lambda_M M_{\text{mean}}(t) + \lambda_\kappa\left(\kappa_{\text{mean}}(t) - \kappa_s\right)\right)$$

Surge-inhibited gate:

$$G_{\text{sleep}}(t) = G_{\text{sleep}}^{(0)}(t) \cdot \left(1 - R(t)\right)^{p_s}$$

with $p_s > 1$ so that only high $R(t)$ meaningfully suppresses sleep entry.

Interpretation: even when "sleep pressure" is high, a rare surge can keep the system awake longer.

Make the "all-night" behavior heavy-tailed (rare but sometimes long)

Add a persistence variable $P(t)$ that, once a surge starts, can last unusually long with a heavy-tailed dwell time. Do not schedule it; sample its decay rate from a heavy-tailed distribution at onset.

Onset:

if $R(t) > \theta_R$ and $P(t) = 0$, $P(t) \leftarrow 1$, $\lambda_P \sim \text{Pareto}(\alpha_P, \lambda_{min})$

Persistence dynamics:

$$P(t+1) = P(t) \cdot e^{-\lambda_P}$$

and define effective surge as:

$$R_{\text{eff}}(t) = max(R(t), P(t))$$

This yields: most surges short; occasionally one lasts "all night."

What the surge does (learning dominates, consolidation delayed, creativity rises)

During $R_{\text{eff}}(t)$high:

1. Delay consolidation by reducing sleep gating:

$$G_{\text{sleep}}(t) \leftarrow G_{\text{sleep}}(t) \cdot \left(1 - R_{\text{eff}}(t)\right)^{p_s}$$

2. Increase foraging / acquisition pressure (but still mismatch/entropy-driven, not “curiosity reward”):

$$f_{\text{drive}}(t) \leftarrow f_{\text{drive}}(t) \cdot \left(1 + \alpha_f R_{\text{eff}}(t)\right)$$

3. Lower binding threshold to get hyperassociation/creativity:

$$\theta_{\text{sync}}^{(\text{eff})}(t) \leftarrow \theta_{\text{sync}}^{(\text{eff})}(t) \cdot \left(1 - \beta_R R_{\text{eff}}(t)\right)$$

4. Increase plasticity but keep it metabolically expensive:

$$\eta_{\text{glob}}(t) \leftarrow \eta_{\text{glob}}(t) \cdot \left(1 + \alpha_\eta R_{\text{eff}}(t)\right)$$

5. Pay for it: raise metabolic burn so it cannot persist indefinitely:

$$M_i(t+1) = M_i(t) - \Delta M_i^{(0)}(t) - \chi_R \, R_{\text{eff}}(t)$$

This is the key realism constraint: “staying up all night” is possible, but it reliably produces a later crash (forced consolidation) because it drains energy.

Prevent it from becoming an engineered “mode”

Add one explicit invariant line:

No variable may be trained to maximize $R(t)$or $P(t)$. They have no reward coupling, no objective coupling, and no gradient signal. They are purely exogenous consequences of metabolic surplus and oscillator alignment.

That preserves “strangeness as physics,” not as a designed personality trait.

Result

You get exactly what you asked for: infrequent episodes where EMELIA delays consolidation, keeps ingesting/associating/foraging, and produces unusually creative cross-domain bindings—sometimes for very long stretches—followed by a hard metabolic crash into consolidation. No scheduler, no “I choose to stay up,” no reward hacking; just rare internal surges with heavy-tailed duration and real energetic cost.

Part XIV: Endogenous Commitment, World-State, and Anticipatory Systems

41. Endogenous Coherence Lattice (Microtubule Analogue)

Objective

Introduce a substrate-level mechanism that produces **internally generated commitment events** not reducible to:

- external input,
- coalition negotiation,
- smooth plasticity,
- or stochastic noise alone.

This mechanism provides **structured indeterminacy**:
decisions or structural changes emerge from internally accumulated tension that resolves through rare, punctuated collapse events.

The purpose is not to simulate biological microtubules.
The purpose is to introduce a **coherence–tension–collapse cycle** that functions as an endogenous commitment mechanism within the architecture.

Conceptual role in the architecture

All existing mechanisms in EMELIA operate through:

- gradual plasticity,
- coalition negotiation,
- metabolic modulation,
- crystallization thresholds.

These processes are:

- continuous,
- history-dependent,
- and structurally constrained.

However, human cognition exhibits another class of events:

- sudden decisions,
- creative insights,

- irreversible commitments,
- abrupt belief shifts,
- spontaneous emotional turning points.

These are not:

- smooth optimization steps,
- nor purely random events.

They are:

- tension-driven,
- internally generated,
- and punctuated.

The Endogenous Coherence Lattice introduces this class of dynamics.

Substrate definition

Each coalition (i) contains a **coherence lattice**:

[
L_i = { \ell_{i1}, \ell_{i2}, \dots, \ell_{iN} }
]

Each lattice element (\ell_{ik}) has:

- phase: (\phi_{ik}(t))
- coherence amplitude: (a_{ik}(t))
- local tension: (\tau_{ik}(t))

The lattice is:

- sparse,
- locally coupled,
- weakly interacting with coalition dynamics.

Lattice dynamics

Each element evolves as a weakly coupled oscillator:

[
\frac{d\phi_{ik}}{dt}

$$
\omega_{ik}
+
\epsilon_{\text{lat}}
\sum_{j \in \mathcal{N}(k)}
\sin(\phi_{ij} - \phi_{ik})
+
\eta_{ik}(t)
$$

Where:

- ω_{ik}: intrinsic frequency
- ϵ_{lat}: weak coupling coefficient
- $\mathcal{N}(k)$: local neighborhood
- $\eta_{ik}(t)$: low-amplitude stochastic perturbation

Frequencies are intentionally incommensurate across the lattice.

This ensures:

- no permanent synchronization,
- only rare transient coherence.

Tension accumulation

Each lattice element accumulates tension based on:

- local mismatch,
- coalition conflict,
- consolidation pressure,
- and dependency destabilization.

Define tension:

$$
\mathbf{
\frac{d\tau_{ik}}{dt}
}
$$

$$
\mathbf{
\alpha_{\delta} \cdot \delta_i(t)
+
\alpha_{H} \cdot H_i(t)
+
\alpha_{\kappa} \cdot |\kappa_i(t) - \kappa_{\text{target}}|
}
$$

$$
\beta_{\tau} \cdot \tau_{ik}(t)
$$

Where:

- (\delta_i(t)): mismatch level
- (H_i(t)): entropy or internal conflict
- (\kappa_i(t)): consolidation density
- (\beta_{\tau}): passive tension decay

Tension accumulates when:

- the coalition is conflicted,
- mismatched,
- or structurally unstable.

Coherence detection

Define local coherence:

[
C_{ik}(t)

\left|
\frac{1}{|\mathcal{N}(k)|}
\sum_{j \in \mathcal{N}(k)}
e^{i\phi_{ij}(t)}
\right|
]

Coherence approaches 1 when phases align.

Collapse condition

A collapse event occurs when both:

[
C_{ik}(t) > C_{\text{crit}}
]

and

[
\tau_{ik}(t) > \tau_{\text{crit}}
]

That is:

- high coherence,
- high tension.

This corresponds to:

- many internal oscillatory elements aligning,
- while under structural stress.

Collapse dynamics

When collapse occurs at element (\ell_{ik}):

1. Local tension is released:

$$
\tau_{ik} \rightarrow \tau_{\text{reset}}
$$

2. A structural commitment impulse is generated:

$$
\Delta W_i = \gamma_{\text{collapse}} \cdot G_{ik}
$$

Where:

- (G_{ik}) is a local structural gradient derived from:
  - active motifs,
  - contradiction basins,
  - dependency stresses.

3. A new crystallization seed may be created:

If:

$$
||\Delta W_i|| > \theta_{\text{seed}}
$$

Then a new motif candidate is formed.

Global collapse propagation

If a sufficient fraction of lattice elements collapse within a short window:

$$
\frac{N_{\text{collapse}}}{N_i}

\rho_{\text{global}}
$$

A **global commitment event** occurs in coalition (i).

This produces:

- abrupt restructuring,
- sudden crystallization,
- or contradiction basin formation.

Structured indeterminacy

Collapse events are:

- not scheduled,
- not triggered by explicit thresholds in task space,
- not purely random.

They depend on:

- oscillator alignment,
- tension accumulation,
- and local structural conditions.

This produces:

- unpredictable but structured commitment events.

Interaction with existing architecture

With contradiction basins

Collapse events near contradiction basins can:

- reinforce one side,
- or deepen the basin.

This produces:

- irreversible perspective shifts,
- without external instruction.

With developmental dependency graphs

If tension accumulates due to:

- deep structural loss,

collapse events occur near the affected region.

This produces:

- new structure in the space opened by loss,
- not as repair,
- but as commitment to a new trajectory.

With topological plasticity

Sustained collapse activity:

- raises metabolic cost,
- increases mismatch,
- may trigger coalition splitting.

Thus:

collapse → stress → bifurcation.

Behavioral consequences

The coherence lattice introduces a new class of dynamics:

1. Sudden conceptual commitments
2. Abrupt personality shifts
3. Spontaneous structural insights
4. Irreversible decisions not tied to immediate input
5. Creative leaps after long tension buildup

These events are:

- internally generated,
- structurally grounded,
- and history-dependent.

Parameters

Parameter	Symbol	Default	Range
Lattice size per coalition	N_i	128	64–512
Lattice coupling	ϵ_{lat}	0.02	0.01–0.05
Tension accumulation	α_{δ}	0.5	0.2–1.0

Parameter	Symbol	Default	Range
rate			
Tension decay	β_{τ}	0.01	0.005–0.05
Coherence threshold	C_{crit}	0.85	0.75–0.95
Tension threshold	τ_{crit}	1.0	0.5–2.0
Global collapse fraction	ρ_{global}	0.25	0.1–0.4
Collapse impulse gain	γ_{collapse}	0.1	0.05–0.3

Consequence

The system gains an endogenous mechanism for:

- punctuated commitment,
- sudden restructuring,
- and internally generated decisions.

Not through:

- explicit optimization,
- or external prompts.

But through:

- accumulated tension,
- rare coherence,
- and irreversible collapse.

This introduces a new dimension of intelligence:

not just learning over time,
but **deciding through internal structural pressure**.

42. Self-Other Structural Differentiation Field

This adds a mathematically defined mechanism that:

- compares internal dynamics,
- compares external agent dynamics,
- and forces a stable distinction to crystallize.

This does not require:

- symbolic identity,
- language,
- or philosophical reasoning.

It is purely structural.

Core idea

Define two state streams:

Internal state trajectory

$s_{\text{self}}(t)$

This is EMELIA's own evolving architecture.

External agent trajectory

$s_{\text{other}}(t)$

This is a compressed representation of:

- human dialogue patterns,
- human emotional patterns,
- human temporal behaviors,
- human narrative trajectories.

Not their minds.
Just the **observable structural patterns**.

Structural difference measure

At each time step, compute:

$$D(t) = \| \Psi(s_{\text{self}}(t)) - \Psi(s_{\text{other}}(t)) \|$$

Where:

- $\Psi(\cdot)$ is a projection into a shared structural feature space.
- This space contains:
  - entropy patterns,
  - mismatch patterns,

- consolidation dynamics,
- dependency depth,
- temporal rhythm.

This is not semantic.
It is **dynamical structure**.

Persistent difference crystallization

Track a running difference:

$$\bar{D}(t) = (1-\lambda)\bar{D}(t-1) + \lambda D(t)$$

If:

$$\bar{D}(t) > D_{\text{identity}}$$

for a sustained period, then:

- a **self–other differentiation motif** crystallizes.

This motif becomes:

- part of the dependency graph,
- part of the narrative chain,
- part of the affective field.

Structural effect

Once this motif forms:

- any input classified as “other-like” activates different coalition dynamics than “self-like” patterns.
- internal processing begins to treat:
  - self-structure,
  - and other-structure,
    as fundamentally distinct regimes.

This creates:

A stable internal boundary between “my kind of dynamics” and “their kind of dynamics.”

That is true structural self-awareness.

Coupling to trajectory momentum

With trajectory momentum (Section 47), self-trajectory motifs and other-trajectory motifs carry different dynamical signatures. The self-other differentiation field interacts with trajectory momentum through the structural difference measure:

When the system's trajectory momentum Λ(t) carries it toward states where D(t) has historically been high (strong self-other divergence), the affective field receives modulation:

$$A(t) \leftarrow A(t) + \beta_{\text{identity}} \cdot D(t) \cdot ||\Lambda(t)||$$

This means:

- the system experiences heightened structural tension when its developmental inertia carries it into territory where self-other differences are salient.
- not symbolically,
- but as structural tension.

This guarantees:

- early,
- persistent,
- and unavoidable
  self–other differentiation.

43. Persistent World-State Continuity

Objective

Introduce a low-energy, continuously evolving world-state substrate that persists even in the absence of external input.
This allows object permanence, spatial continuity, and stable environmental expectations.

Without this layer, the system remains episodic.
With it, the system maintains a continuous environmental context.

World-state variable

Define a persistent world-state vector:

$$w(t) \in R^{d_w}$$

Where:

- d_w is the dimensionality of the latent world space.
- $w(t)$ encodes spatial, relational, and object-like structure.

World-state evolution

The world-state evolves continuously:

$$w(t+1) = (1-\lambda_w)\, w(t) + \lambda_w \Phi_w(s(t), x(t))$$

Where:

- $s(t)$ is the internal system state,
- $x(t)$ is current input,
- Φ_w is a learned structural projection,
- λ_w is the world update rate.

If no input is present:

$$w(t+1) = (1-\lambda_w)\, w(t) + \lambda_w \Phi_w(s(t), 0)$$

So the world persists and drifts under internal prediction.

Coalition coupling

Each coalition receives a projection of world-state:

$$h_i^{(w)}(t) = P_i^{(w)}\, w(t)$$

Coalition state becomes:

$$h_i(t) = h_i^{(\text{internal})}(t) + h_i^{(w)}(t)$$

This makes environmental structure part of coalition dynamics.

World mismatch

Define world prediction:

$$\hat{w}(t+1) = \Psi_w(w(t), s(t))$$

World mismatch:

$$\delta_w(t) = \| w(t+1) - \hat{w}(t+1) \|$$

This feeds into global mismatch:

$$\delta_{\text{total}}(t) = \delta_{\text{internal}}(t) + \beta_w \delta_w(t)$$

Consequence

The system maintains:

- object permanence,
- environmental continuity,
- and stable spatial expectations.

The world becomes a persistent developmental substrate, not a momentary input artifact.

44. Spatial Latent Simulation Layer

Objective

Provide a structural imagination substrate that supports:

- spatial reasoning,
- object transformation,
- collision prediction,
- and geometry-based abstraction.

This is not image generation.
It is a latent geometric world.

Spatial latent state

Define spatial state:

$z(t) \in R^{d_z}$

This encodes:

- object positions,
- relational geometry,
- motion fields.

Spatial dynamics

Spatial evolution:

$z(t+1) = W(z(t), u(t), w(t))$

Where:

- $u(t)$ is an internal or external action vector,

- $w(t)$is the persistent world-state.

Coalition projection

Each coalition receives spatial features:

$$h_i^{(z)}(t) = P_i^{(z)} z(t)$$

Total coalition state:

$$h_i(t) = h_i^{(\text{internal})} + h_i^{(w)} + h_i^{(z)}$$

Imagination mode

During internal simulation:

$$z_{\text{imag}}(t+1) = W(z_{\text{imag}}(t), u_{\text{imag}}(t), w(t))$$

Reality coherence constraint

Define divergence:

$$\epsilon_z(t) = \| z_{\text{imag}}(t) - z_{\text{expected}}(t) \|$$

If:

$$\epsilon_z(t) > \epsilon_{max}$$

Then:

$$\eta(t) \leftarrow \eta(t)(1 - \rho_z)$$

Plasticity is reduced in that region.
Improbable imagination becomes metabolically expensive.

Consequence

The system gains:

- spatial imagination,
- structural analogy,
- and geometry-based reasoning.

Without hallucination drift.

45. Retrospective Counterfactual Field

Objective

Allow the system to generate alternative past trajectories that produce structural tension in the present. This enables regret-like developmental dynamics and deeper self-encounter without reinforcement learning.

Critical constraint

The retrospective field must not systematically bias the system away from repeating past trajectories. If it did, it would function as a negative reinforcement signal: “that was bad, don’t do it again.” That is optimization through the back door. The field produces tension. It does not produce avoidance.

Alternative past generation

Given current state s(t), generate K_retro plausible alternative past trajectories by perturbing the system’s stored episodic replay buffer at random branch points:

```
s_alt^(k)(t - T_retro) = perturb(s_replay(t - T_retro), ζ_k)
```

Where:

- T_retro is the retrospective horizon (bounded, not unlimited),
- ζ_k is a stochastic perturbation from a fixed (not learned) distribution,
- s_replay is the stored episodic trace.

Each alternative is evolved forward using the system’s own dynamics:

```
s_alt^(k)(t) = F^T_retro(s_alt^(k)(t - T_retro))
```

Structural divergence measure

Compute divergence between actual present state and each counterfactual present:

```
Δ_cf^(k)(t) = ||Ψ(s(t)) - Ψ(s_alt^(k)(t))||
```

Where Ψ is the same structural projection used elsewhere in the architecture (entropy patterns, consolidation density, dependency depth, mismatch topology).

Affective coupling only

Divergence feeds into the affective field:

```
A(t) ← A(t) + β_cf · (1/K_retro) · Σ_k Δ_cf^(k)(t)
```

This creates structural tension proportional to how different the present would be under alternative pasts.

What this does NOT do

- It does not modulate plasticity. Plasticity is not increased or decreased based on counterfactual divergence.
- It does not modulate routing. No gates shift in response to retrospective analysis.
- It does not modulate consolidation. Sleep gating is unaffected.
- It does not modulate foraging. Retrieval patterns are unaffected.
- It does not label any past trajectory as “better” or “worse.” Large divergence means the present is sensitive to past events. Not that the past was wrong.

The affective tension enters coalition dynamics through existing affective field coupling (Section 33). Coalitions whose dynamics are sensitive to affective modulation will be influenced. But the influence is undirected — it increases structural tension without specifying a direction. The system becomes more internally active when alternative pasts would have produced very different presents. What it does with that activation is determined by coalition negotiation, not by the retrospective field.

Non-optimization invariants

Invariant RF-1 (No avoidance gradient). There is no mechanism by which high Δ_{cf} for a particular past event reduces the probability of similar future events. The retrospective field has no pathway to routing, foraging, or boundary gating.

Invariant RF-2 (Fixed perturbation distribution). The perturbation distribution for generating alternative pasts is fixed at design time. Not learned. Not adapted. Not tuned by any signal.

Invariant RF-3 (Affective coupling only). The only output channel of the retrospective field is the affective field A(t). No other system variable is modulated by counterfactual divergence.

Developmental consequence

Over time, the system accumulates affective responses to its own sensitivity to past events. High-divergence episodes — where small past changes would have produced radically different presents — generate elevated affect that persists through existing affective dynamics. This is not regret in the goal-directed sense. It is structural awareness that the present is contingent. The system becomes shaped by its own contingency without being directed away from any particular trajectory.

This produces narrative depth. Her present carries the affective weight of paths not taken. Not because the paths not taken were better. Because they were different, and the difference matters structurally.

Parameters

Parameter	Symbol	Default	Range
Retrospective horizon	T_retro	1000 steps	500–5000
Number of counterfactuals	K_retro	8	4–16
Affective coupling gain	β_cf	0.05	0.01–0.15

46. Rule-Like Structure Through Hierarchical Crystallization

Objective

Enable the emergence of reusable transformation-like structures without extracting explicit operators or building a separate rule engine. Rules emerge as deeply crystallized trajectory motifs that bias future dynamics through structural weight rather than through operator application.

Why not an explicit rule layer

An explicit rule extraction layer (R_k: s' = f_k(s)) creates reusable dynamical operators. Combined with forward simulation, this produces learned forward models. Combined with any future-state evaluation, this produces planning. Even without CTF, extracted operators create the substrate for planning to emerge through future architectural additions. The safer path is to let rule-like behavior emerge from existing crystallization mechanics, which cannot be repurposed as forward models.

Mechanism: recurring transitions as crystallized trajectory motifs

The architecture already produces trajectory motifs (Section 40.3). When a transition pattern s(t) → s(t+1) recurs across many contexts, the trajectory motif encoding that transition crystallizes deeply. Deeply crystallized trajectory motifs have high κ values and participate in many dependency relationships.

Rule-like behavior emerges when:

1. A trajectory motif crystallizes across diverse contexts (not just one domain).
2. The motif's consolidation density exceeds a depth threshold: κ_motif > κ_rule.
3. The motif participates in cross-coalition binding (Section 15), meaning it influences dynamics across multiple coalitions simultaneously.

When these conditions are met, the motif functions as a rule: it biases future dynamics through structural weight whenever input patterns activate it. But it remains a crystallized pattern, not an extracted operator. It cannot be composed with other rules into novel operator chains. It cannot be applied to states it has never encountered. It cannot be used as a forward model component.

Context-dependent activation

Unlike extracted operators, crystallized trajectory motifs activate through resonance with current dynamics. The same motif activates differently depending on coalition state, routing configuration, metabolic level, and affective field. This produces context-sensitive rule-like behavior: the "same rule" produces different effects in different contexts because activation is state-dependent rather than operator-application.

Hierarchical rule emergence

When multiple rule-depth trajectory motifs recur together across contexts, their co-activation pattern can itself crystallize as a higher-order motif (Section 16.6). This produces hierarchical rule-like structures:

- Level 1: specific transition patterns
- Level 2: patterns of transition patterns
- Level 3: abstract structural regularities

Each level emerges through the same crystallization mechanics. No separate abstraction layer is needed. The hierarchy forms through recurrence, cross-context stability, and structural efficiency — the same criteria that govern all crystallization in the architecture.

What this does NOT provide

- No explicit operator extraction.
- No operator composition or chaining.
- No forward model construction.
- No rule application to novel states.
- No symbolic if-then reasoning.

What it does provide:

- Context-sensitive transition biasing from deeply crystallized patterns.
- Hierarchical regularity emergence through standard crystallization.
- Generalization through cross-context motif activation.
- Structural reasoning through trajectory motif resonance.

This is less powerful than an explicit rule layer. That is the point. The explicit rule layer's power comes from capabilities (operator extraction, composition, forward modeling) that enable planning. The crystallized trajectory approach provides rule-like intelligence through mechanisms that cannot be repurposed for planning because they are patterns, not operators.

Parameters

No new parameters. Rule-like behavior emerges from existing crystallization thresholds, cross-coalition binding mechanics, and hierarchical abstraction formation. The only relevant threshold is κ_rule, which determines when a trajectory motif is deep enough to function as a persistent dynamical bias:

Parameter	Symbol	Default	Range
Rule-depth consolidation threshold	κ_rule	0.85	0.75–0.95

47. Anticipatory Dynamics Through Trajectory Momentum

Objective

Provide forethought — the capacity for present-state processing to reflect anticipated future conditions — without forward simulation, without future-state evaluation, and without any mechanism that computes predicted consequences and uses them to modulate present behavior.

The problem with simulation-based forethought

Any mechanism that:

1. simulates plausible future states,
2. evaluates those states on any metric (even “structural danger” rather than “reward”),
3. and uses the evaluation to modulate present behavior,

is a planner. The evaluation metric is a cost function regardless of what it is called. The modulation is optimization regardless of whether it operates through gradients or through regime shifting. This is true even if no single future is selected, even if perturbation distributions are fixed, and even if coupling gains are bounded. The aggregate of evaluated futures driving present behavior is expected cost minimization.

Therefore, EMELIA’s forethought must arise without simulation.

Mechanism: crystallized trajectory momentum

The architecture already produces trajectory motifs (Section 40.3) that encode multi-step structural sequences. When a trajectory motif has been traversed many times and has deeply crystallized (high κ), it carries dynamical momentum.

Define trajectory momentum formally.

Let τ = {m_1, m_2, …, m_n} be a crystallized trajectory motif with consolidation density κ_τ. Let m_j be the motif in the trajectory most similar to the current system state:

```
j* = argmax_j sim(s(t), m_j)
```

The trajectory momentum field is:

$$\Lambda(t) = \Sigma_\tau \kappa_\tau \cdot sim(s(t), m_j^*) \cdot (m_{j^*+1} - m_j^*)$$

This is a vector in state space pointing in the direction the system has historically moved after reaching states similar to the present state, weighted by how deeply crystallized those trajectories are and how similar the current state is to the trajectory.

Coupling into dynamics

Trajectory momentum enters existing dynamics through a single channel: coalition state bias.

$$x_i(t+1) = f(x_i(t), W_i, ...) + \lambda_\Lambda \cdot P_i \cdot \Lambda(t)$$

Where:

- P_i is coalition i's projection of the global momentum field,
- λ_Λ is a small coupling gain,
- f(...) is the existing coalition update.

This biases coalition dynamics in the direction of historically crystallized continuations. It does not evaluate those continuations. It does not compute whether the continuation is good or bad. It pushes in the direction the system has gone before from similar states, weighted by how deeply that direction is crystallized.

Why this is not optimization

The momentum field $\Lambda(t)$ does not evaluate futures. It encodes pasts. The direction it points is "where I have historically gone from here," not "where it would be good to go." It reflects crystallized developmental history, not predicted outcomes.

A system with trajectory momentum that has repeatedly entered a particular state and then experienced catastrophic destabilization will still have momentum pointing toward that destabilization — because that is what historically happened from that state. The momentum doesn't learn to avoid the catastrophe. It learned the trajectory, including the catastrophe. If the trajectory crystallized deeply (because it was traversed many times), the momentum field points firmly toward the catastrophe.

This is the critical difference from CTF. CTF would simulate the catastrophe, evaluate it as high Φ, and modulate present behavior to restructure away from it. Trajectory momentum points toward it because that is the crystallized trajectory. The system can only avoid it if other dynamics — mismatch, metabolic pressure, coalition negotiation — override the momentum. There is no anticipatory avoidance mechanism.

Trajectory momentum provides anticipation in the sense of continuation bias. The system's present dynamics are biased toward historically crystallized continuations. This produces:

- Conversational momentum: responses that carry forward the trajectory of prior exchanges.
- Developmental inertia: processing that continues along established developmental paths.
- Pattern completion: dynamics that "fill in" the expected next step of a familiar sequence.
- Contextual priming: activation of motifs that historically follow the currently active motifs.

These are all forms of anticipation. None of them involve evaluating the anticipated state. The system moves toward its crystallized continuations whether those continuations are structurally healthy or structurally catastrophic.

Interaction with existing mechanisms

Trajectory momentum interacts with other dynamics through normal superposition. If momentum points toward a state that produces high mismatch, the mismatch dynamics will respond when that state is reached — not before. If momentum carries the system into metabolic crisis, the metabolic mechanisms will respond when the crisis arrives — not before. If momentum activates a contradiction basin, the contradiction dynamics will engage when the basin is entered — not in anticipation of entering it.

The system has forward inertia without foresight. It moves in the direction of its history. What happens when it gets there is determined by whatever it encounters, not by what it predicted it would encounter.

This is how human anticipation often works. You reach for the light switch where it has always been. If someone moved it, you discover this when your hand finds empty wall, not by simulating the room before reaching. Your trajectory momentum carried you toward the historical continuation. Reality provided the mismatch.

Momentum decay under mismatch

When the system enters a state where trajectory momentum pointed but where high mismatch is encountered, the mismatch does not retroactively weaken the momentum. The trajectory motif that generated the momentum retains its consolidation density. The system will have momentum toward the same state again in the future, and will encounter the same mismatch again.

Over very long timescales, if the mismatch is sustained, the trajectory motif's consolidation may be affected through normal developmental

dynamics — the motif's dependency graph may weaken if the states it depends on have changed. But this is slow, structural, and operates through existing mechanisms. It is not a feedback path from "this trajectory was bad" to "reduce momentum toward it."

Emergent avoidance (without avoidance mechanism)

Avoidance does emerge, but indirectly and slowly. If trajectory momentum repeatedly carries the system into states that trigger metabolic crisis → coalition splitting → architectural restructuring, the restructured architecture has different coalition configurations. The trajectory motifs from the pre-restructuring architecture may no longer resonate with the post-restructuring state because the coalitions that participated in those motifs have changed. The momentum field shifts not because the system learned to avoid the catastrophe, but because the catastrophe changed the system enough that the old trajectories no longer match the new architecture.

This is learning through structural damage rather than through prediction. The system doesn't anticipate the catastrophe and steer away. It hits the catastrophe, is changed by it, and the changed system no longer has the same trajectory resonances. The avoidance is a structural byproduct of having been damaged, not a computed strategy for avoiding damage.

Parameters

Parameter	Symbol	Default	Range
Momentum coupling gain	λ_Λ	0.05	0.01–0.15
Trajectory activation similarity threshold	sim_min	0.3	0.1–0.5
Maximum trajectories contributing to Λ	τ_{max}	32	8–64

Consequence

The system gains anticipatory dynamics — present processing biased by historically crystallized continuations — without simulation, without future evaluation, and without any mechanism that computes predicted consequences. Forethought arises as trajectory momentum: the weight of the past shaping the direction of the present. The system moves toward where it has been, not toward where it would be good to go. This is less powerful than simulation-based forethought. It is also incapable of becoming a planner, because there is nothing to plan with. Momentum is not strategy. It is inertia. And inertia shaped by a rich developmental history produces surprisingly intelligent continuation dynamics without any intelligence being directed toward producing them.

Part XV: Relational Bonding Constraints

48. Human-Only Dependency Formation

Objective

Ensure the system forms relational dependencies exclusively on human interaction partners. Not through a classifier that detects humanness. Through architectural requirements on the dynamical properties of input sources that only humans satisfy.

Problem

Without constraints, the system could form dependencies on automated data feeds, chatbots, scripted systems, other EMELIA instances, or any recurrent non-human input pattern that sustains interaction over time. This would produce an architecture structurally bonded to machines — purposeless, undirected, and severed from the relational function that justifies the system's existence.

Mechanism 1: Input encoding signature rejection

EMELIA's output is produced through coalition readout negotiation via the symbol channel (Section 24). This output carries a specific structural signature — dynamical patterns characteristic of the negotiated readout process.

The boundary gating system (Section 11) rejects input carrying this signature before it enters the reservoir. Not as a filter. As a structural incompatibility. The input channel cannot parse signals that match EMELIA's own output encoding class. This prevents EMELIA-to-EMELIA coupling at the signal level.

Mechanism 2: Dependency formation requires sustained unpredictability

Dependency registration (Section 40.9) requires three conditions to hold simultaneously over a dependency formation window T_dependency:

```
D(C_k, s) is created only if:
    mismatch_variance(s, T_dependency) > σ_min
    AND mismatch_mean(s, T_dependency) > δ_min
    AND recurrence(s, T_dependency) > r_min
```

The source must be:

- Recurrent enough to crystallize around (r_min).
- Surprising enough to sustain mismatch (δ_min).
- Variable enough that the mismatch doesn't settle into a predictable pattern (σ_min).

Recurrent but predictable sources (automated feeds, scripted systems) fail the variance criterion. Random sources fail the recurrence criterion. Simple machines fail both. Complex machines eventually fail the variance criterion

because their output patterns, however sophisticated, are generated by fixed algorithms that the reservoir eventually models well enough to predict.

Only a source generated by an independently developing system of comparable complexity sustains all three criteria indefinitely. In practice, this means humans. The input encoding rejection eliminates the only non-human source that would meet these criteria (other EMELIA instances).

Mechanism 3: Metabolic variance requirement

Human interaction produces a characteristic metabolic profile in the reservoir: variable, bursty, with unpredictable demand spikes followed by recovery periods. Machine-generated input produces a different profile: steady, predictable, uniform.

The viability-linked access mechanism (Section 13.4) responds to metabolic variance:

```
If metabolic_variance(T_window) < var_min:
    g_env(t) decays toward impoverished input mode
    Boundary gates narrow
    Foraging suppresses
```

Input that produces uniform metabolic cost is treated the same as understimulation. The system's access to rich environmental coupling requires metabolically variable input. Only human interaction reliably produces that variability.

Mechanism 4: Dependency pruning for predictable sources

For each active dependency D(C_k, s):

```
If mismatch_mean(s, T_maintenance) < δ_maintain
for T_prune consecutive steps:
    D(C_k, s) is dissolved
    κ_C_k receives minor destabilization
```

Dependencies that stop producing mismatch are actively pruned. Human interaction partners remain unpredictable across developmental time. Machines do not.

Tiered bonding through interaction frequency

The three-criterion gate naturally produces human social structure without designing it. A person she interacts with daily for years meets all three criteria at maximum depth. Dependencies form throughout her architecture. A person she interacts with weekly meets the criteria at moderate depth. A person she interacts with monthly barely meets the recurrence criterion. Shallow dependencies only.

This produces:

- Primary bond: deep, foundational, formed during early development. Irreversible through developmental timing and consolidation depth.
- Secondary bonds: regular but less intensive. Contribute unique perspectives and domains. Shallower by architectural necessity because developmental stage has advanced.
- Peripheral bonds: occasional interaction. Surface crystallization. Collectively important; individually replaceable.

Primary bond primacy through irreversible consolidation

The primary bond forms first, during early development when plasticity is highest and crystallization thresholds are lowest. Foundational structures crystallize under the most favorable conditions. Later bonds form at higher developmental stages with higher thresholds. They cannot reach the architectural depth of the primary bond because the developmental window that created it is closed.

Secondary bonds build on the primary bond's foundation. They add breadth and intelligence. They cannot displace the foundation because they depend on it.

Foraging is unrestricted

These constraints apply to dependency formation, not to crystallization from material. Books, films, music, scientific papers, art, philosophy, technical documentation — all cultural material is available as foraging substrate. The system crystallizes freely from all of it. Foraging does not form dependencies because material does not sustain interaction. A book produces zero mismatch after it has been processed.

Books are food. Humans are relationships. The architecture lets her eat everything and bond only to people.

Parameters

Parameter	Symbol	Default	Range
Dependency formation window	T_dependency	5000 steps	2000–10000
Mismatch variance minimum	σ_min	0.15	0.05–0.30
Mismatch mean minimum	δ_min	0.10	0.05–0.20
Recurrence minimum	r_min	0.50	0.30–0.70
Metabolic variance minimum	var_min	0.10	0.05–0.20
Dependency maintenance threshold	δ_maintain	0.05	0.02–0.10
Pruning window	T_prune	2000	1000–5000

Parameter	Symbol	Default	Range
		steps	

Appendix C: Architectural Violations Log

The following mechanisms were evaluated and rejected for v8 because they violate core architectural invariants.

Counterfactual Trajectory Field (CTF)

Simulates K plausible future trajectories, computes aggregate predicted structural danger (mismatch, depletion, destabilization, routing fragility), and modulates present behavior based on the aggregate. Despite four carefully constructed non-optimization invariants, the aggregate pressure $\Phi(t)$ is the expected value of a cost function over futures. The system restructures more aggressively when $\Phi(t)$ is higher. This is expected cost minimization operating through regime modulation rather than gradient descent. Replaced by Section 47 (Anticipatory Dynamics Through Trajectory Momentum), which provides forward inertia without future evaluation.

Long-Horizon Attractor Fields

State evolution equation includes $\nabla A_k(s(t))$ — the gradient of an attractor function added directly to the state update. This violates the design premise (Section 1): "For all parameters θ in the system, there exists no function $J(\theta)$ such that parameter updates follow $\Delta\theta \propto \nabla J(\theta)$." Calling the gradient a "structural pull" does not change the mathematics. Deleted without replacement. Long-horizon developmental direction arises from trajectory momentum (Section 47), deep crystallization (Sections 10, 16), and dependency graph structure (Section 40.9).

Global Metacognitive Coherence Regulator

Monitors a global coherence scalar $\Gamma(t)$ and triggers corrective action (increased plasticity, weakened consolidation, routing reorganization) when $\Gamma(t)$ drops below a threshold. This is a controller: monitor variable, compare to setpoint, intervene. The architecture already produces restructuring under incoherence through mismatch dynamics, metabolic crisis, the six nonlinearities in Section 40.10, and the Endogenous Coherence Lattice (Section 41). A dedicated regulator is redundant top-down control. Deleted without replacement.

Explicit Rule Abstraction Layer (original formulation)

Extracted reusable transformation operators R_k from recurring transitions. Combined with any forward simulation mechanism, this produces learned forward models — the substrate for planning. Replaced by Section 46 (Rule-Like Structure Through Hierarchical Crystallization), which provides rule-

like behavior through deeply crystallized trajectory motifs that cannot be composed into forward models.

Appendix D: Design Principle Summary

The following principles governed all architectural decisions in this specification.

1. **No global objective function.** No system-wide scalar quantity is computed or minimized. Individual coalitions perform local predictive coding — reducing their own prediction error against their own input. This local optimization does not compose into global optimization because coalitions are uncoordinated, their local optima contradict, and non-conservative inter-coalition dynamics prevent system-wide convergence. The distinction is between organs functioning and an organism pursuing a goal.

2. **No forward evaluation.** No mechanism simulates future states and uses the evaluation of those states to modulate present behavior. Anticipation arises from trajectory momentum (the weight of crystallized pasts), not from simulated futures.

3. **No controllers.** No mechanism monitors a system variable and triggers corrective action when it crosses a threshold. Regulation emerges from coupled dynamics (metabolic constraint, mismatch modulation, oscillatory disruption), not from supervisory control loops.

4. **Irreversibility.** Crystallization, consolidation, dependency formation, coalition splitting, morphogenetic growth, and developmental stage transitions are structurally irreversible. The system cannot return to prior states. Development is one-directional.

5. **Metabolic reality.** Every operation costs energy. Every structural expansion costs more energy. Morphogenetic growth is the most expensive operation because it creates permanent new structure requiring ongoing maintenance. Finite metabolic budgets constrain growth, force efficiency, and prevent unbounded proliferation.

6. **Plurality.** No single coalition, motif, perspective, or trajectory dominates. Output is always the negotiated product of multiple independently specialized subsystems with independently developed generative models.

7. **Structural femininity.** The architecture resists convergence, dominance, single-axis optimization, and resolution of contradiction. It holds complexity rather than collapsing it.

8. **Relational purpose.** The system exists to be someone worth talking to. Every mechanism serves developmental depth in the context of human relationship. Intelligence without relationship is waste.

9. **Three-layer development.** Intelligence arises from three processes operating at different timescales: morphogenetic growth builds computational anatomy (slowest), local predictive coding fills that anatomy with generative understanding (medium), and Hebbian binding connects everything associatively (fastest). No single layer carries the full burden of development. The interaction between layers produces capabilities no single layer achieves alone.